# Impact of Interruption Frequency on Nurses' Performance, Satisfaction, and Cognition During Patient-Controlled Analgesia Use in the Simulated Setting

2015

Kristi Campoe
*University of Central Florida*

# IMPACT OF INTERRUPTION FREQUENCY ON NURSES' PERFORMANCE, SATISFACTION, AND COGNITION DURING PATIENT-CONTROLLED ANALGESIA USE IN THE SIMULATED SETTING

by

KRISTI R. CAMPOE
MSN, Nova Southeastern University, 2007
BSN, University of Texas Medical Branch, 1995

A dissertation presented in partial fulfillment of the requirements
for the Degree of Doctor of Philosophy in Nursing
in the College of Nursing
at the University of Central Florida
Orlando, Florida

Summer Term
2015

Major Professor: Steven Talbert

**ABSTRACT**

Interruption during medication administration is a significant patient safety concern within health care, especially during the administration of high risk medications in nursing practice. Specifically, patient-controlled analgesia (PCA) devices are frequently associated with adverse events and have a four-fold increased risk of patient injury compared to non-PCA related adverse events. While the nature and frequency of interruptions have been established for nurses' medication processes, the impact of interruption frequency on nurses' PCA interaction has not been fully described or measured. The purpose of this study was two-fold: (a) to quantify the impact of interruption frequency on registered nurses' performance, satisfaction, and subjective workload during PCA interaction, and (b) determine nurses' perceptions of the impact of interruption frequency.

This study employed a mixed-method design. First, an experimental repeated measures design was used to quantify the impact of interruption frequency. Nine registered nurses (RN) were recruited from Florida hospitals. The RNs completed PCA programming tasks in a simulated laboratory nursing environment for each of four conditions where interruption frequency was pre-determined. Established human factors usability measures were completed for each of the four test conditions. RN performance was video-recorded with time-stamp then analyzed for performance measures of efficiency (total task time) and effectiveness (accuracy). RNs completed a user satisfaction survey and subjective workload assessment (NASA-TLX). The research questions were answered using repeated measures analysis of variance with (RM-ANOVA), McNamar's test, and Friedman's test. After each experiment, semi-structured interviews were used to collect data that were analyzed using inductive qualitative content analysis to determine nurses' perceptions of the impact of interruption frequency.

The sample of RNs ($n$=9) was female (100%) working full-time in the medical-surgical setting. Total time to complete tasks (seconds) ranged from 189.00 to 419.00 ($M$=292.11, $SD$=73.25). For accuracy, total number of errors for participants ranged from 0 to 6 ($M$=1.56, $SD$=94.71). Five (56%) participants reported a low impact of interruption frequency on their satisfaction. Subjective workload scores for the NASA-TLX (raw) for condition A ($M$=23, $SD$=10.87) and condition B ($M$=26.00, $SD$=11.14) ranged from 12.00 to 47.00; the range was highest for condition C ($M$=31.56, $SD$ 22.31) at 12.00 to 78.00. The research questions were answered using repeated measures analysis of variance with (RM-ANOVA), McNamar's test, and Friedman's test. Results of the RM-ANOVA were significant for the main effect of interruption frequency on efficiency $F(3,24)=9.592$, $p = .000$. McNemar's test did not show significance for the impact of interruption frequency on effectiveness (accuracy). Friedman test showed participant satisfaction was significantly impacted by interruption frequency ($x^2$=9.47, $df$=3, $p$=0.024). Friedman test showed no significance for the main effect of interruption frequency on subjective workload scores by condition type ($x^2$=1.88, $df$=3, $p$=0.599). Results of the qualitative content analysis revealed two main categories to describe nurses' perception of interruption frequency: the *nature of interruptions* and *nurses' reaction to the interrupted work environment*.

The results suggested that interruption frequency significantly affected efficiency (task completion time) and satisfaction for participants but not participant effectiveness (accuracy) or subjective workload scores. The high error rate during PCA programming tasks indicated the need to evaluate the conditions in which nurses complete PCA programming as each error is potential risk of patient harm or injury. Interruption frequency may lead to time pressure that negatively impacts total task time and accuracy. Nurses' described the impact of interruption

frequency as having a negative impact on the work environment and subsequently implement compensating strategies to counterbalance the impact of interruption in the workplace. Nurses perceive that patient safety is negatively impacted by frequent interruption and nurses experience negative intrapersonal consequences as a results of frequent interruption, that have the potential to negatively impact performance, satisfaction, and subjective workload. Additional study is needed to better understand the impact of interruption frequency on nurses' performance effectiveness (accuracy) and subjective workload.

## ACKNOWLEDGMENTS

This dissertation is dedicated to my loving and supportive husband, Alex. Your love sustains me. Thank you.

This has been a long and fruitful journey. I have learned much about myself and my capacity as a human, student, nurse, researcher, and educator. Completion of dissertation research is not possible without the support of family, friends, mentors, and advisors.

Thank you to Lucas, Lauren, Vanessa, and Nicholas for your unending patience, love, encouragement, and support. It is not every day that children of your age are in a position to support a parent year-after-year. I could never have accomplished this lifelong goal without your love each day. I hope to inspire you to achieve your highest potential, to pursue life-long learning, and to accomplish excellence in the career of your choice. I love each of you!

Mom and Dad, I offer my thanks to you for a lifetime of support and encouragement. Thank you for everything you do!

Thank you to my dearest friends, especially from soccer and gymnastics who have supported me with kind words and helped with the all four kids, helping to get them to school, games, practices, and meets. There are so many people to thank but I want to specifically off my gratitude to Tim and Tina M., Shamon and Yolanda H., Patrick and Andrea M., Mike and Linda C., Jeff and Sabrina H., Kim S., and Laura P.

Thank you to my fellow doctoral students who supported me throughout coursework, long study hours, and during the dissertation process. You gave me just the support I needed, right when I needed it! I offer my sincere appreciation and thanks to Angela R., Sylvia T., Janet H., Jenny C., Vicki S., Rich R., Sherry A., and Valerie L.

Thank you for your support and mentoring my academic and professional endeavors, especially Dr. Diane Whitehead, Dr. George Smith, Risë Sandrowitz, Dr. Doris Parish, Dr. Karen Giuliano, Dr. Carina Vocisano-Bruel, Dr. Sally Spooner, and for my dear friend and mentor, Wanda Dearth.

In memorial, I offer a special thanks to two academic mentors who passed away during my doctoral program. Thank you to Dr. John S. Barnett who opened my eyes to the study of human factors in healthcare and nursing. I offer my sincerest appreciation and heartfelt respect for my mentor and advisor Dr. Jacqueline Byers. She convinced me to commit to the UCF PhD program and guided my doctoral coursework and early dissertation topic planning. Without Drs. Barnett and Byers and their support, I would not have found my way. Thank you.

I would like to thank my dissertation committee for their advice and support throughout the dissertation process from proposal through completion of my dissertation study. Thank you Dr. Mary Lou Sole, Dr. Diane Andrews, and Dr. Florian Jentsch. Thank you to Dr. Steven Talbert for serving as my dissertation chair. I am grateful to and thankful for the faculty at the UCF College of Nursing for the excellent teaching and support during my coursework. Special thank you to Dr. Anne Norris for her teaching, as well as presentation and publication opportunities.

Finally, I would like to thank the Tampa Bay Organization for Nurse Executives (TBONE) for a doctoral scholarship. I would like to thank Nova Southeastern University (NSU) for financial support via the President's Faculty Research and Development Grant (PFRDG). TBONE and NSU made my dissertation research financially possible.

**TABLE OF CONTENTS**

## LIST OF FIGURES

## LIST OF TABLES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| df | degree of freedom |
| FDA | U.S. Food and Drug Administration |
| HF/E | Human Factors and Ergonomics |
| IOM | Institute of Medicine |
| ISO | International Organization of Standards |
| LTM | Long-term Memory |
| M/S | Medical/Surgical |
| MAE | Medical Administration Error |
| NASA-TLX | National Aeronautical and Space Administration-Task Load Index |
| p | probability |
| PCA | Patient-controlled analgesia |
| RM-ANOVA | Repeated Measures-Analysis of Variance |
| RN | Registered Nurse |
| SMCIMD | System Model of Clinician Interaction with Medical Devices |
| WM | Working Memory |
| $x^2$ | Chi-square |

## CHAPTER ONE: INTRODUCTION

The seminal Institute of Medicine (IOM) reports that launched recent programs to improve healthcare quality and safety called for the adoption of technology as one approach to address quality and safety issues in health care (IOM, 2001; Kohn, Corrigan, & Donaldson, 1999). Subsequent technological advances led to the adoption and refinement of medical devices and systems that automate certain high-risk patient care processes. Despite the promise of technology to improve medication administration and patient safety, human error continues to exist and unintended consequences of technology compromise nursing practice and patient safety.

Medication errors in the acute care setting are an unintended consequence of patient care and they occur at any point during the four phases of the medication process: (a) provider prescription by physicians and advanced practice nurses, (b) transcription and verification of medication orders usually by clerical workers or computer systems, (c) dispensing and delivery to patient care areas by pharmacy staff, and (d) administration of the medication to the patient primarily by nurses (Leape et al., 1995). Medication administration errors (MAEs), errors in the final phase of the process, are a unique concern to nursing practice in that nurses bear the major responsibility for medication administration and as they often include the use of complex medical devices for medication administration and are least likely to be intercepted by safety structure or processes (Bates et al., 1995). MAEs are the most common adverse drug event (Wong et al., 2009) and occur as frequently as one MAE per patient per day (Lin & Ma, 2009). The severity MAEs range from insignificant delays in administration of medications that are not time sensitive to lethal overdoses of powerful parenteral medications. The outcomes of MAEs are 1.5 million injured patients yearly with treatment costs from drug-related injury in a hospital

setting approximated to be $3.5 billion yearly in the United States (Aspden, Wolcott, Bootman, & Cronenwett, 2007).

New studies continue to test the error reducing benefits of healthcare systems and devices (Buntin, Burke, Hoaglin, & Blumenthal, 2011; Maddox, Danello, Williams, & Fields, 2008) while other studies identify that systems and devices that are poorly designed leading to new classes or types of errors such as interface errors, content errors (Kushniruk, Triola, Borycki, Stein, & Kannry, 2005; Zhang, Patel, Johnson, Chung, & Turley, 2005) and, poor feedback to users (Obradovich & Woods, 1996). These studies affirm our limited understanding of user interaction with complex systems and devices and related medication administration errors. A clearer understanding of the nurse-device interaction is needed to reduce medication administration error rates involving complex medical devices, like patient-controlled analgesia.

## Introduction to the Problem

Interruption during medication administration is a significant patient safety concern, especially during the administration of high risk medications involving patient-controlled analgesia (PCA). PCA-related errors are more harmful and costly than non-PCA-related errors. Current evidence describes the nature and frequency of interruptions during nurses' medication administration processes, but the impact of frequency and intensity of interruption during nurses' PCA interaction has not been addressed in the literature. Improved understanding of nurses' interaction with PCA could reduce PCA-related errors and improve patient safety.

## Background of the Study

Medical device safety has become a leading patient safety concern, and safety data link serious injury and death to patient-controlled analgesia (PCA) technology used in healthcare

facilities. Although PCA is an effective interactive process for the administration of narcotic analgesia, allowing patients to have control over their pain management while eliminating delays in administration (Hudcova, McNicol, Quah, Lau, & Carr, 2011), PCA is known to be associated with frequent adverse events. Specifically, PCA is delivered using a complex automated infusion device, which has been shown to have a four-fold increased risk of patient injury when compared to non-PCA related adverse events (Hicks, Sikirica, Nelson, Schein, & Cousins, 2008; Schein, Hicks, Nelson, Sikirica, & Doyle, 2009). Two thousand four hundred ninety-seven PCA related adverse events were documented between 2003 and 2004 (Meissner et al., 2009) and 9,571 events were reported between 2000 to 2005 (Hicks et al., 2008). The calculated mean cost of PCA-related errors resulting in patient injury was $6,943, as compared to $28 for a PCA-related error without patient injury (Meissner et al., 2009). The relative high risk, frequency, and cost of PCA-related errors support the need to identify contributing factors and in the long-term, develop strategies to minimize risk, mitigate error, and decrease cost of with these serious errors.

Distractions and interruptions occur frequently in healthcare (Redding & Robinson, 2009; Rivera-Rodriguez & Karsh, 2010) and negatively affect the human-device interaction such as those involving PCAs (Swayze & Rich, 2011). Nurses reported that interruptions during medication administration rounds create a higher risk for error (Palese, Sartor, Costaperaria, & Bresadola, 2009). Observation of interruption frequency during nurses' medication administration found that up to two-thirds of medication administration rounds were interrupted (Kreckler, Catchpole, Bottomley, Handa, & McCulloch, 2008; Palese et al., 2009). Nurses' perceptions regarding the impact of interruption frequency and interruption intensity during complex medical device use, specifically PCA use, have not been described in current literature.

However, interruption frequency observed during nurses' medication administration tasks ranged from 0.8 to 41.8 events per hour (Biron, Loiselle, & Lavoie-Tremblay, 2009).

The human-device interaction occurs between four components: user, device, tasks (e.g., PCA pump programming), and the environment of use. Each component influences the outcomes of the human-device interaction within a complex work-system. Although some medical device usability studies have described interruptions during infusion pump programming and device use (Lin et al., 1998; Lin et al., 2001), they typically have not reported the nurses' subjective experiences of interruptions during PCA use. Qualitative methods have been used to describe usability problems within the context of system interaction (Garmer et al., 2002a; Rose et al., 2005; Staggers, 2003; Staggers, Kobus, & Brown, 2007). These studies have not included nurses' perception the impact of interruptions or nurses' perceptions of intensity of interruption during interactions with medical devices such as PCAs.

**Statement of the Problem**

While it is known that interruption is reported to contribute to PCA-related adverse events (Hicks et al., 2008) and that the severity of medication errors increases with interruption frequency (Westbrook, Woods, Rob, Dunsmuir, & Day, 2010), the impact of interruption frequency during nurses' PCA-interaction is not known. Currently, no empirical study exists to quantify the effects of interruption frequency on nurses' PCA-interactions or determine nurses' perceptions of the impact of interruption frequency and interruption intensity.

**Purpose of the Study**

The purpose of this study is twofold: (1) to quantify the impact of interruption frequency on registered nurses' performance, satisfaction, and subjective workload during PCA interaction and (2) determine nurses' perceptions of the impact of interruption frequency.

**Research Design**

A mixed-methods approach will be used for this research. First, an experimental repeated measure crossover design will be used to quantify the impact of interruption frequency for aims one and two of the study. After each experiment, semi-structured interviews will be used to collect data that will be analyzed to determine nurses' perceptions of the impact of interruption frequency and interruption intensity on their PCA interactions for aim three of the study.

**Study Aims, Research Questions, and Hypotheses**

This specific aims of this study are listed below. The central hypothesis of this study was that interruption frequency during nurses' patient-controlled analgesia device interaction will affect nurses' performance efficiency and effectiveness, subjective satisfaction, and perceived subjective workload.

*Aim #1: Determine the impact of interruption frequency on nurses' PCA performance.*

- *Research question #1:* What is the effect of interruption frequency on the efficiency and effectiveness of medical-surgical nurses' PCA use?
- *Hypothesis #1:* Increased frequency of interruption will have a negative effect on nurses' performance efficiency (EF1-task completion time) and effectiveness (A1-accuracy).

*Aim #2: Determine the impact of interruption frequency after PCA interactions on medical-surgical nurses' subjective satisfaction and subjective workload.*

- *Research question #2:* What is the effect of interruption frequency on medical-surgical nurses' subjective satisfaction and subjective workload with PCA use?
- *Hypothesis #2.* Increased interruption frequency will decrease nurses' subjective satisfaction and increase subjective workload.

*Aim #3: Determine nurses' perceptions of the impact of interruption frequency on nurses' PCA interactions.*

- *Research question #3*: What are medical-surgical nurses' perceptions of the impact of interruption frequency during PCA interactions?

**Conceptual Framework**

This study proposes a conceptual framework (Figure 1) to evaluate the human-device interactions using closed system model of system inputs, interaction processes, and outputs (Campoe, 2013b). System components interact, dependent upon human capability and limitations, resulting in PCA-related medication adverse events involving nurses. First, consistent with usability methods (International Organization of Standards [ISO], 1998, 2007) there are four system input factors: user (nurse), PCA device, PCA programming tasks, and interruption frequency environment. Next, process factors are human limitations and abilities of human cognition and attention that impact and help explain human-device interaction (Wickens & Holland. 2000; Wickens & McCarley, 2008). Finally, output factors will be measures of efficiency, effectiveness, satisfaction, and subjective workload (ISO, 1998, 2007) which aid in the determination for system changes. This framework assumes that the output of the interaction is context-dependent to the environment where devices are used and that change in any one system component impacts outcomes.

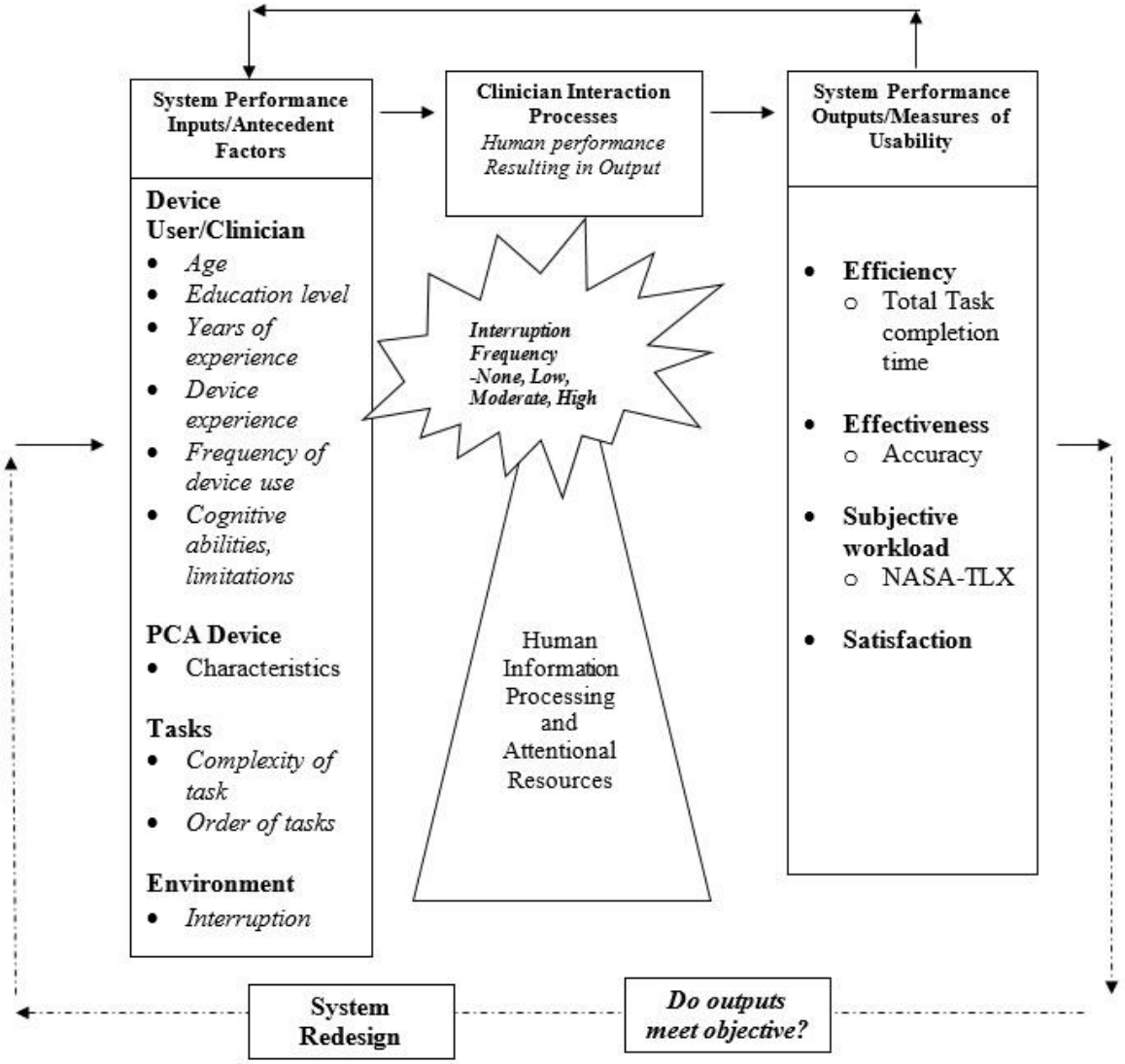



Figure 1. System model of clinician interaction with medical devices (SMCIMD) (Campoe, 2013b).

## Definition of Terms

The following terms and definitions are used in this dissertation:

- *Interruption* is the human experience that creates discontinuity in task performance situated within a specific context (Brixey et al., 2007).
- *Interruption frequency* is the rate of auditory or visual stimuli perceived by a nurse.

- *Efficiency* is the time and human resources consumed in order to complete tasks (ISO, 1998; Hornbæk, 2006).
- *Effectiveness* is the level of accuracy at which users achieve specified tasks (ISO, 1998; Hornbæk, 2006).
- *Satisfaction* is the user attitude toward the use of a device, system, or product (ISO, 1998; Hornbæk, 2006). An attitude a settled way of thinking or believing about someone or something (Oxford Dictionaries, 2015).
- *Subjective workload* is the human mental or cognitive effort expended during human-device interactions (ISO, 1998; Hornbæk, 2006).

**Significance of the Study**

The U.S. Food and Drug Administration [FDA] (2010; 2011) and other agencies (e.g., Benjamin, 2008; Institute for Safe Medication Practices, 2010) have major initiatives in progress to reduce medication adverse events and improve patient safety. This proposed study specifically supports these efforts. This study will contribute to existing patient safety research by quantifying the impact of interruption on the nurse PCA interaction, and will lay the groundwork for future study in this area.

For nursing, the use of a human factors systems approach and measures are novel to the study of interruption during nurses' PCA interactions. This proposed study builds upon existing knowledge regarding nurses' perceptions of medication administration specifically contributing new knowledge regarding nurses' perceptions of interruption frequency and intensity during PCA interactions. New knowledge can be used by manufacturers to improved human factors

PCA infusion device design. Organizational and individual –based interventions may be necessary to mitigate unnecessary interruption.

## Conclusion

Studies reporting the frequency and nature of interruption do not consider the use of complex medical devices that are frequently incorporated by nurses during medication administration such as infusion pumps or patient controlled analgesia (PCA) systems. For example, when nurses are interrupted during PCA set-up, it is not known if interruptions impact proper materials and device set-up including programming tasks. Further, it is not known if interruption impacts the accuracy and timeliness of the tasks, or cognitive resources to safely interact with the PCA. Determining nurses' subjective experience of interacting with patient PCA systems will build upon existing knowledge. Knowledge regarding nurses' perceptions of interruptions and their characteristic intensity during PCA interaction may improve our understanding PCA-related errors and support development of interventions to improve in patient safety.

## Outline for the Remainder of the Dissertation

Chapter Two includes a review of the literature, to be followed by the Methodology in Chapter Three, Results in Chapter Four, and the Discussion and Conclusions in Chapter Five.

## CHAPTER TWO: LITERATURE REVIEW

Patient-controlled analgesia (PCA), a method of pain control designed to allow the patient to administer preset doses of analgesic whenever the patient sees fit, has been determined to be an effective interactive process for the administration of narcotic analgesia (Crisp et al., 2012). PCA allows patients to control their own pain management, while eliminating delays in administration (Hudcova, McNicol, Quah, Lau & Carr, 2011); however, enormous safety concerns related to the administration of PCA have been identified. Many of these concerns are related to opioids, and numerous cases have pointed to human error, specifically during the ordering, dispensing, or administering of PCA (Hicks, Sikirica, Nelson, Schein, & Cousins, 2008). The frequency of PCA-related adverse events from 2003 to 2004 alone numbered 2,497 (Meissner et al., 2009), and 9,571 events were reported from 2000 to 2005 (Hicks et al., 2008). The calculated mean cost of PCA-related errors resulting in patient injury was $6,943, as compared to $28 for a PCA-related error without patient injury (Meissner et al., 2009).

Hospital staff members, registered nurses, anesthesia providers, prescribing personnel, dispensing personnel, and administering personnel assume the responsibility of quality care to all hospital patients, which necessitated this researcher to explore the dynamics of human error and stimuli that may alter staff members' attention given to their tasks at hand. External stimuli related to nurses interacting with multiple machines and persons in a working environment shift attention away from tasks related to patient-controlled analgesia (Hicks et al., 2008).

This literature review explores current research literature focused on foundational theories that analyzed the relationship between humans and machines, medical device usability, and human performance factors. The review of the literature begins with the conceptual foundation for a modern theoretical model that measures the interactive relationship between

humans and machines. Three conventional conceptual and theoretical models were utilized to construct a new, separate conceptual system model, which was designed and implemented for this study. The established conceptual and theoretical models that were examined were (a) the ISO model of usability (ISO, 1998), which focused on the device usability; (b) the human-machine interaction (Czaja, 1997; Shackel, 1991) model, which focused on the relationship between human and machine; and (c) human information processing theory, which strictly focused on human ability (Wickens & Hollands, 2000; Wickens, 1992). Because the models used in previous research failed to explain how context of use variables impact human performance and satisfaction, the System Model of Clinician Interaction with Medical Devices (SMCIMD), a synthesized model of the aforementioned models, created a unique and new model to be implemented in this study.

Following the section about the conceptual foundation of the SMCIMD, studies were analyzed that focused on system performance input, nurse interaction processes, and system performance output in relation to the conceptual framework of the SMCIMD. The purpose of this review is to present an analysis and synthesis of published medical device usability studies and related literature. The application of human-factors usability methods improved medical device use and safety throughout the product-life cycle (Braun, 2005; Shah & Robinson, 2007). However, it has been noted that relatively few medical device usability analyses have been published in the peer-reviewed literature (Fairbanks, Caplan, Bishop, Marks, & Shah, 2007; Martin, Norris, Murphy, & Crowe, 2008).

Finally, this review of literature details an overview of current medical device usability, the limitations of current studies that examined the link between the usability of a medical device, and the frequency of interruption that device caused in relation to a nurse's subjective

workload and tasks. Usability was established only after user abilities and limitations have been described and considered within the design and use of a medical device. Studies in this review show usability was limited by level of nursing experience (Garmer, Liljegren, Osvalder, & Dahlman, 2002; Ginsburg, 2005), and that device-related experience may not have transferred to new or comparable devices (Nemeth, Nunnally, Bitan, Nunnally, & Cook, 2009; Nunnally & Bitan, 2006). Also, stimuli received and interpreted by device users limited efficiency and effectiveness of medical device use (Carayon et al., 2007; Ginsburg, 2005; Lin, Vicente, & Doyle, 2001). Devices that do not meet user abilities and limitations result in errors, which limit user satisfaction (Lin et al., 2001; Liu, Tech, & Osvalder, 2004), and lead to coping strategies, such as work-arounds and unsafe practices (Brixey, Zhang, Johnson, & Turley, 2009; Carayon et al., 2007; Obradovich & Woods, 1996). The disconnection between user abilities and limitations and device design creates an opportunity for future research.

## Conceptual Foundation

### ISO Model Background and Description

The International Organization for Standardization (ISO) developed the model of usability (Figure 2) to guide manufactures in the description and measure of device usability throughout the device's life-cycle. The ISO model of usability was recognized as an international standard for studying usability across industries. The U. S. Food and Drug Administration (FDA, 2000; 2011) required medical device manufactures to evaluate usability to improve device design, mitigate errors, and improve patient safety. The ISO model described system context of use, measures of usability, and concept relationships to usability goals.

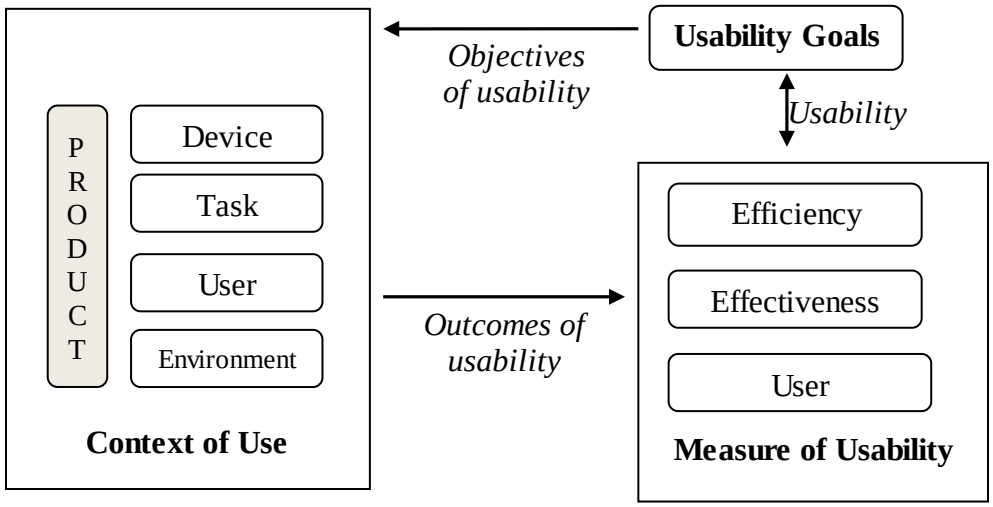



Figure 2. ISO (1998) Model of usability.

**ISO model context of use.** Context of use described the components of the domain interaction, and referred to as the user-device interaction. Characteristics of the users, device or system, tasks, and environment of use are described in detail and should be represented in as realistic manner (ISO, 1998). Moreover, medical device, task, and environmental characteristics were identified as components that worked toward presumed usability goals (Campoe, 2013b).

**ISO model of outcomes of usability.** The ISO model described the distinct outcomes of usability using measurements of effectiveness, efficiency, and satisfaction. Operational definitions were described by ISO (1998), the FDA (2000), and current literature (Hornbæk, 2006). A systematic review of usability measures reported reliability and validity of effectiveness, efficiency, satisfaction, and other measures from the human-computer interaction literature (Hornbæk, 2006). Hornbæk also suggested that the measures focus on macro issues "related to cognitively and socially complex tasks" (p. 98). Challenges were found in some of the

human-computer interaction research literature, which showed scientific testing errors; however, the majority of the research predicted reliable and valid measures. Table 1 summarizes each ISO measure.

Table 1. Measures of usability adapted from ISO 9241-11 (1998) and Hornbæk (2006).

| **Measure of usability** | **Definition** | **Sample measures** |
|---|---|---|
| Effectiveness | Extent to which the intended goals of use are achieved | -Task completion: number or percent of tasks that user successfully completes<br>-Accuracy: accuracy with which user completes task, measured through quantification of error |
| Efficiency | Resources that have to be expended to achieve the intended goals | -Time: duration of tasks or parts of tasks<br>-Mental effort: user physiologic or cognitive/mental effort when using the interface<br>-Interface usage patterns: how user makes use of the interface to solve problems |
| User Satisfaction | Extent to which the user attitude finds the product acceptable | -Measures of user satisfaction, attitude, acceptance, or preference using standardized questionnaires (QUIS) or non-standardized measures such as percent of favorable and unfavorable responses; user choice or rank of preference |

**Usability goals and relationships among variables.** Usability and intended objectives contained a unidirectional relationship with other variables of the model. The ISO model of usability described the context-dependent nature of usability and how to measure usability in terms of effectiveness, efficiency, and satisfaction. The complex interactions between the product or system, user, device, tasks, and environment of use influenced usability within a complex work-system, with a central focus on context of use. Congruent with a systems-perspective, a change to any one of the context variables could result in changes to the outcome variables (Bevan & Macleod, 1994; Shackel, 1991, 2009).

**ISO model usability assumptions.** Being context specific, the validity of usability data was limited to the users, devices, tasks, and environment (ISO, 1998). Usability could be significantly different within different contexts (Czaja, 1997). Next, the effects of change to one component of the system could be measured by user performance (i.e., effectiveness and efficiency) and satisfaction. This assumption was analogous with the human factors system perspective of a system, which was a belief system or a body of principles, methods, and tools focused on the common purpose of a system (Czaja, 1997). With the user at the center of the system, the process of user-centered design (Martin, Norris, Murphy & Crowe, 2008) allowed interventions to achieve established goals and improve the system that supported the human factors system evaluation.

**ISO model utility and evaluation.** Although relatively few medical device studies have been published in peer-reviewed literature, the ISO model and its components were the focus in existing medical device studies. The ISO model has been implicitly used in multiple published, peer-reviewed medical device usability studies (Fairbanks et al., 2007; Garmer et al., 2002; Lin et al., 1998; Lin et al., 2001; Liu et al., 2004; Nunnally & Bitan, 2006; Trbovich, Pinkney, Cafazzo & Easty, 2010). Moreover, the model has been described and applied extensively in usability studies in disciplines including software engineering, aviation and telecommunications (Wicklund, Kendler & Strochlic, 2011). However, it is important to note the limitations of the ISO model.

The ISO model of usability depicted the relationship between usability goals, context of use components, and measures of usability, but has not predicted or explained the effects of changes in one or more context variables on specific outcomes or measures. As a general model, the ISO model has not specifically identified, described, or explained the complex variants that

exist within healthcare and nursing practice environments. Previous medical device usability studies focused on the existence of variability in users, tasks, and environment of use, which affected outcomes of usability (Campoe, 2013a). For example, user experience (Garmer et al., 2002), variability and complexity of tasks (Carayon et al., 2007), and variability and sources of stimuli in the physical environment affected usability (Trbovich, Prakash, Stewart, Trip, & Savage, 2010; Westbrook, Coiera, et al., 2010).

The ISO model of usability has not described human cognitive processes that occurred during the user-device interactions. Human cognition and cognitive processing were important factors in healthcare and nursing, and the research literature reported the use of subjective workload as a measure of efficiency (Hart, 2006; Kataoka, Sasaki & Kanda, 2011; Lin et al., 1998; Lin et al., 2001). To further address this limitation, this review looked beyond the ISO model of usability to integrate cognitive psychology theory.

**Human Cognition in a System: Abilities and Limitations during Device Interaction**

According to Wicklund, Kendler and Strochlic (2011), the strict study of machine usability was flawed in its aim, because the role of human interaction with the machine had to be taken into account. Human physical and cognitive abilities and limitations influenced medical device design to ensure that devices were efficient, effective, and safe (Wicklund et al, 2011). Knowledge of how users interact within a system or with complex devices was central to usability and patient safety. Moreover, humans and machines were integrated within a working environment, and the use of the machine was a human controlled operation.

**The human-machine interaction.** Shackel (1991) described the human-machine system as the relationship between the user, the task, the tool, and the environment. The human-machine

interaction model (Figure 3) (Czaja, 1997) included the Shackel components, and it related the process of interaction between a human user and a system during interaction in the context of use environment. The model details three components of the interaction: the machine-system, human-system interface, and the human.

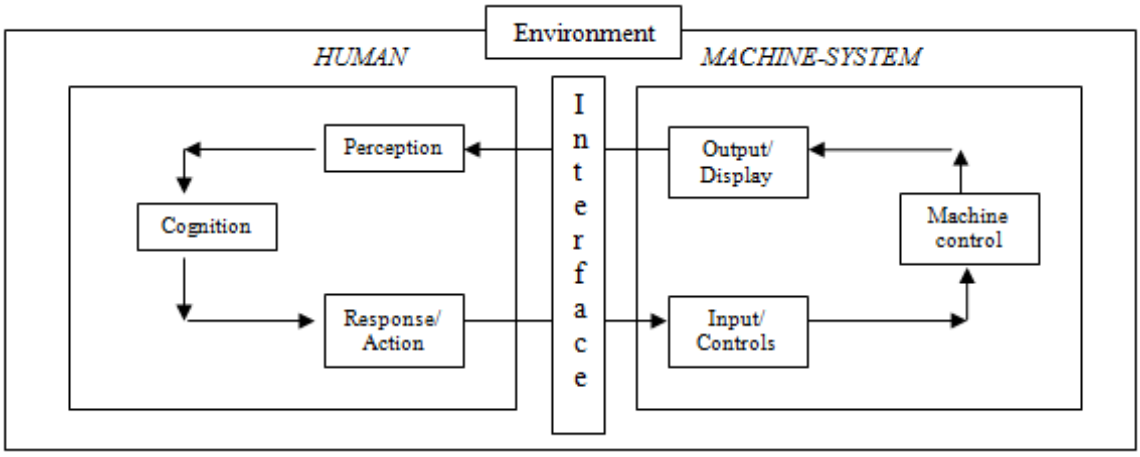



Figure 3. Human–machine model (Czaja, 1997).

The model depicted a closed system that provided a feedback loop as the three components of the interaction were processed. The machine-system had two interaction components. The *output mechanism* was a system component perceived by the user, such as a visual display or auditory alerts that presents information to the user. The input mechanism, a system component such as a keyboard, mouse, voice, dialog boxes, or menu selections, received information from the user. The human-system interface shared information between user and device, during the user-system interaction. The third component was the actual human, the device user, who perceived the information output from the system; information was cognitively perceived and processed. The human response was executed, and the system input mechanism was utilized, completing the interaction cycle.

As expected, normal use occurred within the system environment. The components of the interaction (machine-system, human-system interface, and human), including human performance, was affected by the environment. The model demonstrated that humans reacted to stimuli from both the environment and the machine, while interacting with the machine, with which improvements to the machine interface (i.e. usability) or the environment impact the interaction. Understanding the components and processes of human information processing helped to explain how human performance was impacted by interacting with a system. Exploration of human information processing was necessary to explain how usability and interruptions impact user cognitive processes during medical device use in the workplace environment.

**Human information processing model.** Focused on human thought and cognitive response, the human information processing model (Figure 4) described the many processes throughout an interaction between a human, such as a nurse, and a complex system (Wickens, 1992; Wickens & Hollands, 2000). An abundance of human information processing research was related to human technology use and interaction within a system, including specific research on the user-interface issues with infusing pumps (Schraagen & Verhoeven, 2013). However, the current research was still inconclusive in regards to interruption frequency as the motivating cause of adverse effects in the healthcare field (Grundgeiger & Sanderson, 2009). The gaps in the research literature also extended to the necessity to create a new or modified model to render inclusive cognitive interaction accurately (Langdon, Persad & Clarkson, 2010).

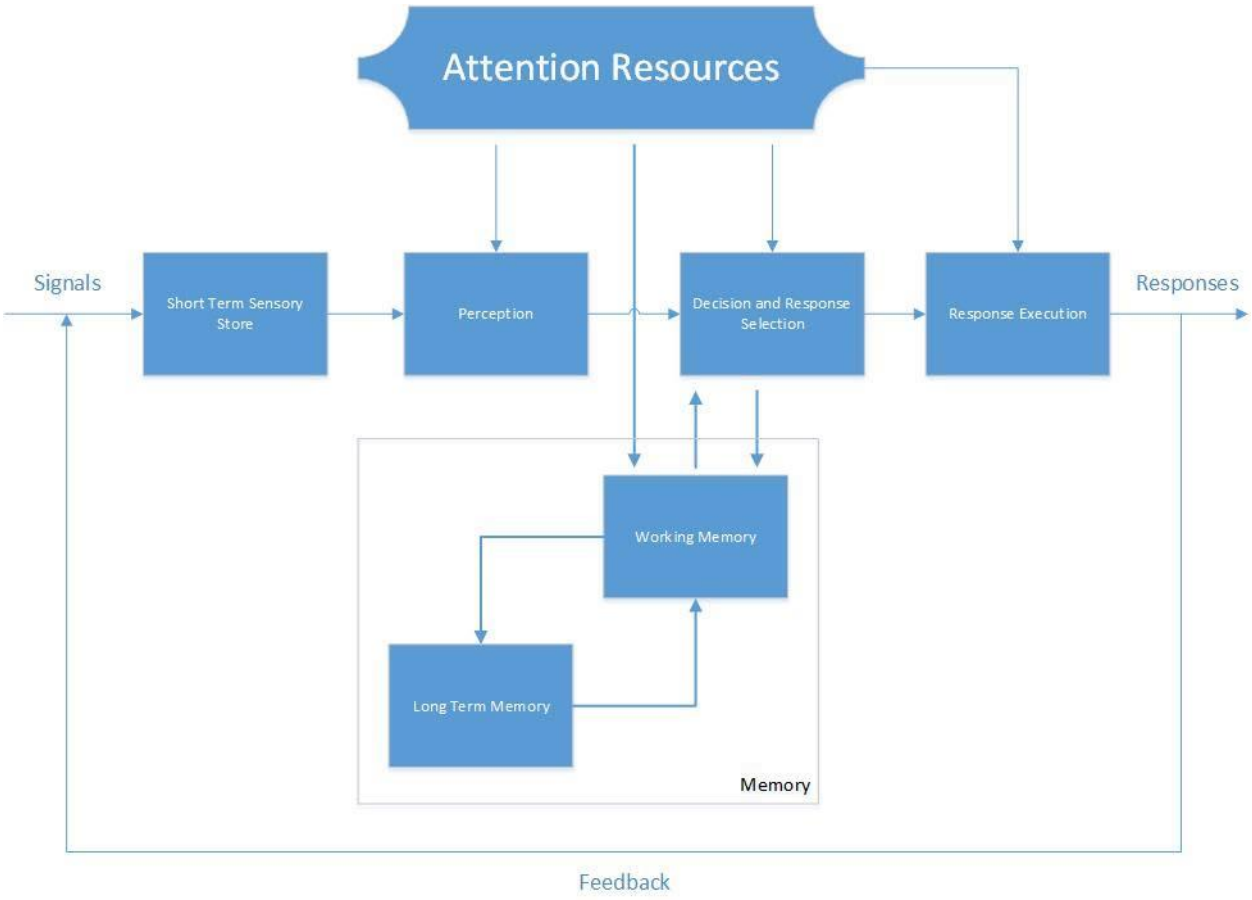



Figure 4. Human information processing model (adapted from Wickens, 1992; Wickens & Holland, 2000).

Stimuli from both the system and the environment were received by the nurse through the senses, and these stimuli were continuously processed. In the model diagram, the feedback loop of human information processing occurred in stages, and the loop was void of a fixed starting point. At any one time, nurses processed internal, such as stress, and external stimuli from the environment, such as auditory and visual stimuli (Potter et al., 2005). In order to gain access to the brain, the sensory memory received and briefly stored stimuli, estimated to be 1/2 second for visual stimuli and 2 to 4 seconds for auditory stimuli (Wickens & McCarley, 2008). The sensory data was transformed for encoding in the Working Memory (WM), which was referred to as

short-term memory in the research literature. Stimuli that were not given meaning or fully encoded by the brain, such as those stimuli that were not perceived, attended to, or interrupted, either eroded or was not transformed for use by the memory (Wickens & McCarley, 2008).

The WM received information, held it for approximately 12 to 30 seconds, and processed all information for preparation in the Long-Term Memory (LTM) (Wickens & McCarley, 2008). Receipt of information for the WM was rapidly and automatically interpreted and given meaning with little attention from the user. Based upon interpretation of sensory data, the nurse either ascribed the meaning of the stimuli in WM (bottom-up processing), or from LTM retrieval of previous experiences (top-down processing). The model suggested that a function of the WM was to retrieve information from the LTM, although the LTM consisted of stored, encoded memories, and life-long unlimited capacity, and WM was used to support recall and recognition perception of new stimuli.

**Human information processing theory utility and evaluation.** Usability studies provided numerous insights into the medical device user abilities and limitations during interactions. Nurses continuously received stimuli through the information processing feedback loop from internal and external sources such as stress, interruption, stimuli in the environment, and interactions with complex medical devices (Hoonakker et al., 2011). Ultimately, nurses experienced multiple sources of stimuli from the naturalistic work environment and experienced high subjective workload (Hoonakker et al., 2011; Kataoka et al., 2011; Potter et al., 2005; Redding & Robinson, 2009; Wolf et al., 2006). Cognitive concepts such as attention, memory, and learning have been incorporated into established usability principles (Nielsen & Mack, 1994; Zhang et al., 2003); therefore, connecting human factors and cognitive science, which lend to a more complete understanding of human interaction with complex devices.

**Limitations of human information processing.** Relevant to nurses' interactions with medical devices, they have developed important limitations with regards to WM and attention in response to interruptions within complex systems. Studies concerning WM suggested that participants were limited to seven pieces of information, plus or minus two (Sörqvist, 2010; Wickens & McCarley, 2008). Only three-to-five items of information could be simultaneously processed.

The multiple resource theory of attention explained that within the context of human information processing, a nurse's attention can be divided between completing complex tasks, such as programming an infusion device requiring auditory and visual processing, while receiving competing auditory or visual stimuli. As an extension of human information processing theory, multiple resources theory (Parasuraman & Manzey, 2010; Wickens & McCarley, 2008) described the capacity of attention resources relevant to nurses' choices to divide attention or allocate attention resources to different tasks or mental processes. Attention was limited by the ability to engage and access the resource within the memory. Potter (2005) suggested that nurses maintain a sustained level of attention while completing nursing processes and care; therefore, the outcome of this process affects subjective workload and performance.

Human information processing and multiple resource theories explained how nurses completed complex tasks, while receiving and processing multiple stimuli. These theories, which will be discussed in a future section of this literature review, could be used to study user-interaction with medical devices or extend the ISO model of usability; however, the aforementioned conceptual and theoretical models were fragmentary when discussing the role of interruption on the nurse's interaction with not only the medical device, but also internal and external stimuli.

**System Model of Clinician Interaction with Medical Devices (SMCIMD)**

A conceptual or theoretical model was needed to communicate the key concepts of a problem for empirical study of nurses' interaction with medical devices and explain or predict the process leading to the problem. No singular existing model supported the study of the effects of interruption frequency during nurses' Patient-Controlled Analgesia (PCA) use in the simulated setting. Accordingly, the conceptual model (Figure 5) for this study was synthesized from the ISO model of usability (Figure 2) (ISO, 1998), human-machine interaction (Figure 3) (Czaja, 1997; Shackel, 1991), and human information processing theory (Figure 4) (Wickens & Hollands, 2000; Wickens, 1992).

This study implemented a conceptual framework (Figure 5) to evaluate nurse-PCA interactions using a closed system model of system performance inputs, nurse interaction processes, and system performance outputs (Campoe, 2013b). The model provided a framework for understanding the effects of interruption frequency during nurses' patient-controlled analgesia (PCA) use and served as a framework for describing nurses' perceptions of interruption frequency during PCA interactions. Developed as a closed system, the model adopted important ISO concepts, and extended the descriptive nature of the ISO model by incorporating the theory of human information processing, as a means to explain the influence of medical device usability and stimuli received and processed by nurses.

The synthesized model drew upon distinct areas from within the Human Factors and Ergonomics (HF/E) literature, along with related theoretical literature, in order to contribute to the development of the SMCIMD. This conceptual model (Figure 5) utilized both the International Organization for Standardization (ISO) model of usability (ISO, 1998) and the human information processing model (Wickens & McCarley, 2008; Wickens & Carswell, 1997).

Literature was also reviewed on cognitive distraction and interruption in healthcare settings, where human information processing was critical to patient safety, and contributed to the model (Figure 5) below. The study developed system model explained the user-experience process by incorporating key concepts into three components: System Performance, Nurse Interaction, and System Outputs.

**System performance inputs.** The four system performance input factors depicted on the left side of the model were developed to remain consistent with the standards set for in the ISO usability model (ISO, 1998, 2007). These input factors provided the context for understanding the elements and characteristics of the system that directly affect the nurse's interaction process. Existing ISO descriptions (Table 1) were incorporated into the SMCIMD conceptual model with consideration of known characteristics that impact medical device usability. The feedback from performance outputs were considered within this model, which affect the performance inputs, and represented a portion of the feedback loop.

**Clinician interaction processes.** Interaction processes were considered to be the abilities and the limitations of human cognition and attention that influence human-device interaction.

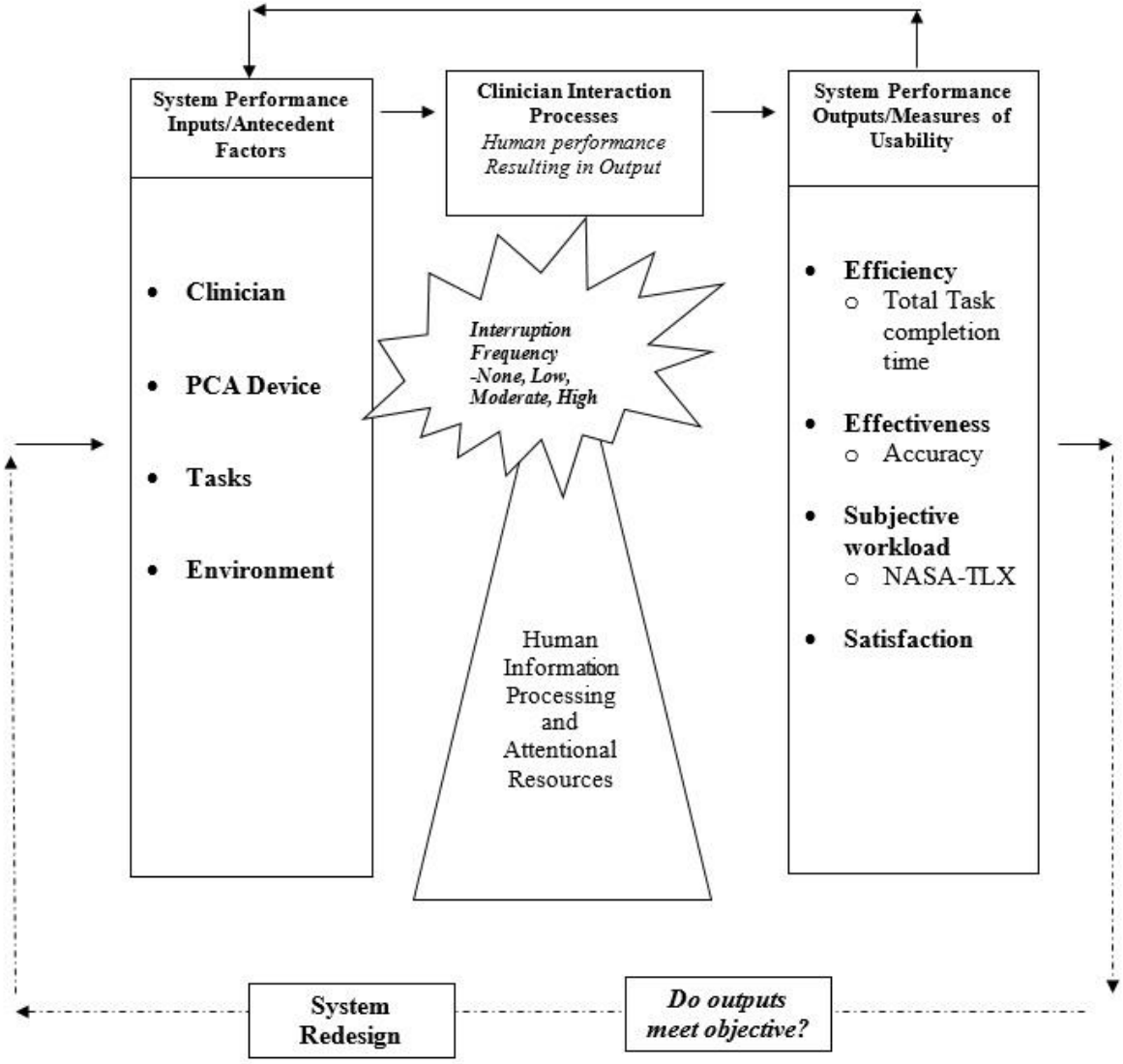



Figure 5. Conceptual framework: System model of clinician interaction with medical devices (SMCIMD).

The nurse interaction processes, the center triangle of the model, represented nurses' interaction with medical devices and the environment. The Human Information Processing Model (Figure 4) explained the cognitive processes nurses use to receive stimuli and the

limitations of attention that were affected by context of medical device usage. Being central to the model, this element developed an important aspect to the model which had been missing in previous studies and research.

**System performance outputs.** Output factors, listed on the right side of the model (ISO, 1998, 2007), aided in the determination for system changes, according to the model. The system components processed by the nurse were dependent upon cognitive capability and limitations, and resulted in measurable human performance. Efficiency, effectiveness, and user satisfaction contributed to the analysis of objectives, which developed system changes necessary to increase performance and completed the feedback loop. This model assumed that the output of the interaction is context-dependent and that change in any one system component impacted the overall outcome. System feedback and redesign formed the feedback loop between system inputs, interaction processes, and performance outputs. In medical device usability, feedback provided valuable information as to whether or not goals were achieved as expected, user performance, and patient safety. Risk mitigation or system re-design should be used when the goals are not achieved.

**Model utility and application.** Incorporating human information processing with the ISO model of usability extended the ISO model into an explanatory model. This was important because it has been established that particular context of use characteristics specifically impacted medical device usability (Campoe, 2013b); however, previous research has not empirically studied the impact of these contexts of use factors in light of human cognitive abilities and limitations. The SMCIMD identified the relationship between context of use variables and outcomes, and measured effectiveness, efficiency, user satisfaction, and subjective workload. Fundamental to the outcome measures, knowledge of human information processing explained

the need for context of use, such as how stimuli from the medical device and the environment are perceived, cognitively processed, and executed by the nurse. For this reason, knowledge of human information processing, an understanding of limitations in memory and attention, and the impact of interruption, were necessary to fully explain human behavior or performance in a complex system. Furthermore, performance measures could be used as a basis of comparison within the same context (Hornbæk, 2006), using an experimental design, and a comparison between the different effects of interruption frequency.

**Nurses' performance of PCA tasks.** System components and interruption factors listed in the SMCIMD affected a nurse's ability to perform patient-tasks, which resulted in adverse PCA-related medication events. Consistent with usability methods (ISO, 1998, 2007), the following four aspects of nurse functions oriented the system input factors: user (nurse), PCA device, PCA programming tasks, and simulated interruption frequency environment. The same four input factors described context of use variables: nurses (users), patient controlled analgesia device (PCA device), programming tasks and sub-tasks (tasks), and interruption frequency (environment, including their intensity). Moreover, the nurses' process factors were the limitations and abilities of human cognition and attention, which impacted human-device interaction (Wickens & Hollands, 2000; Wickens & McCarley, 2008).

Output factors measured efficiency, effectiveness, satisfaction, and subjective workload (ISO, 1998, 2007), which aid in the determination for system changes. Without interruptions from the work environment, and context of use remaining consistent, the dependent variables could be measured while nurses completed programming tasks and sub-tasks with a patient controlled analgesia (PCA) system. This framework assumed that the output of the interaction was context-dependent to the environment where devices were used, and that change in any one

system component impacted outcomes. Understanding nurses' perceptions of the impact of interruption frequency and interruption intensity could be ascertained during or after PCA interaction. This unique nature of the SMCIMD improved understanding of the effects of interruptions based upon nurses' perspectives.

SMCIMD was needed because medical device usability analyses specifically identify error-producing conditions of the user, device, tasks, and environment (Campoe, 2013b). Though studies have measured usability and interaction in terms of efficiency, effectiveness, and satisfaction measures, this study model described how context of use variables impact human performance and satisfaction.

## Medical Device Usability

An integrative review methodology summarized evidence including diverse methodologies within a domain, employing specific strategies to enhance rigor (Whittemore & Knafl, 2005). According the FDA (2011), usability methods should be utilized as complementary and independent approaches to improve design and reduce risks throughout a medical device life-cycle. Also, usability methods should be developed from previously refined usability methods. Figure 6 summarized the search strategy and yields for this study.

The initial search consisted of a keyword search to identify published peer-reviewed English language literature from 1991 to 2013. The search began with 1991 because early technical standards for medical device usability were first presented in draft format. MEDLINE®, Cumulative Index of Nursing and Allied Health Literature (CINAHL®), PsychInfo®, Psych Article®, and Science Direct electronic databases were searched for specific terms to identify potential sources: (1) medical device or medical technology or technology or

device; (2) use or usability or user-computer interaction. The initial search yielded 1,213 sources from various safety-related industries. From that population, the results of the initial screening yielded 253 sources. Next, each source title and abstract were screened in order to identify relevancy to the purpose of this review. The final review resulted in 20 sources that were retrieved, read, and reviewed for purpose and adherence to inclusion/exclusion criteria. Campoe (2013a) published the results of this review in a paper at the 2013 International Symposium on Human Factors and Ergonomics in Health Care.

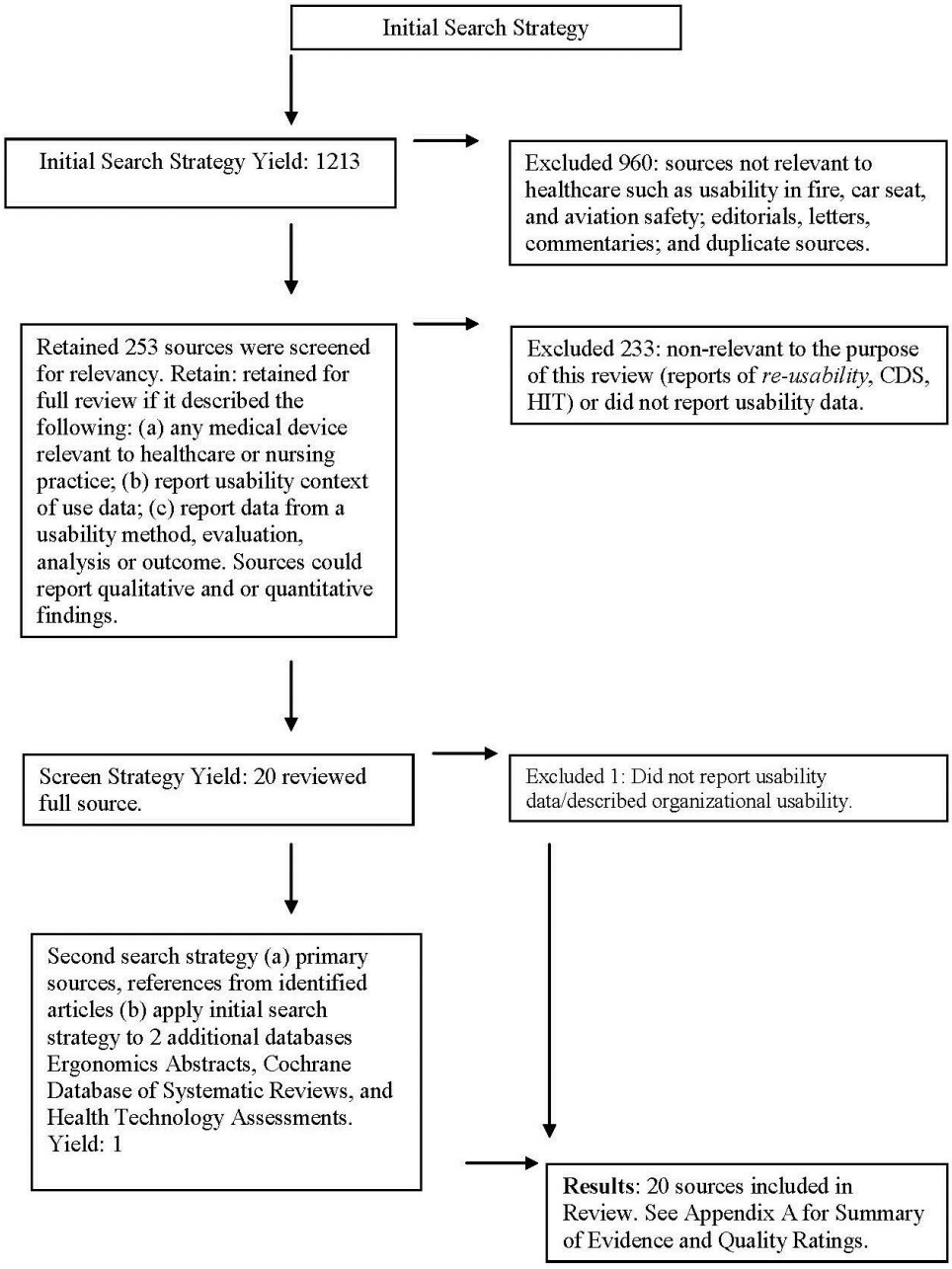



Figure 6. Search strategy and yields.

Design, methods, and research reliability and validity were assessed, and limitations of current studies were identified. Twenty articles were appraised and classified for strength and

quality. Using the hierarchy of evidence for strength of evidence described by Melnyk and Fineout-Overholt (2011), studies were categorized by the strength of evidence into levels one (highest level of evidence) through seven (lowest level of evidence). Six sources were categorized as Level one, ten sources as Level two, and four sources Level three. Sources were ordered, coded, and analyzed to guide clinical practice, education, policy, and future research. From the twenty articles summarized in the evidence table (Appendix A), diverse usability methods and measures were implemented to analyze medical device usability. Fifteen studies utilized multiple, complimentary approaches, while five studies conducted a single-method approach. The literature reviewed will be presented in the following areas according to the study's conceptual framework: system performance input factors, nurse interaction processes, and system performance output factors.

**System Performance Input Factors**

The first section of this review related to the left side of the study model, titled system performance input factors, which were the interactions between the nurse-user, device, task, and environmental components as a system. These were contextual factors that impacted medical device usability. Each study reviewed provided descriptions, and analyzed how each component impacted usability within a system. These descriptions were important because each context component contributed to either safe and effective medical device use, or unsafe, ineffective medical device use (FDA, 2000). Studies reviewed suggested that users and environmental characteristics impact usability.

**Nurse users.** The user experience and level of experience with a medical device impacted usability (Carayon, Hundt & Wetterneck, 2010; Garmer, Liljegren, Osvalder, &

Dahlman, 2002; Ginsburg, 2005). Novices encountered the most usability problems and had difficulty articulating the usability problems (Garmer et al., 2002a). Novices also experienced errors and undetected errors, which limited the efficiency of use (Carayon et al., 2010; Ginsburg, 2005). These findings were consistent with current literature on novice versus expert user-device interactions (Kjeldskov, Skov & Stage, 2010). Novice device users were not experienced enough to recognize or differentiate usability problems from personal performance, abilities, or limitations.

Nonetheless, there was conflicting evidence about the influence of user experience on usability. Nursing students with minimal device-specific experience were more efficient (task time) and more effective (fewer errors) when interacting with a device (Lin et al., 1998) than nurses with extensive device experience who committed high errors rates, when completing tasks on familiar devices (Nemeth, Nunnally, Bitan, Nunnally & Cook, 2009; Nunnally & Bitan, 2006). Also, medical device usability evaluations lost validity when the study employed only experts or users of one experience level (U. S. Food and Drug Administration, 2011).

**Device.** Studies described usability problems as device characteristics that impacted the user experience. Usability problems were relative to device, interface, and system attributes that did not meet user needs, limitations, or abilities, which led to error. Evaluating devices against established design principles was the most common method to identify usability problems (Nielsen & Mack, 1994, Zhang et al., 2003). Table 2 summarizes device analysis and focus, including reviews of hardware, such as switches, buttons, and knobs; software interfaces; or materials, such as training materials, instructions, or tubing cassettes. Eleven studies evaluated the usability of all three aspects. Furthermore, studies that evaluated one aspect of a device may not have identified usability problems within the system (Shah & Robinson, 2007).

Table 2. Medical device descriptions and focus: hardware (H), software interface (I), materials (M).

| Study | Device(s) or set of products description | Focus |
|---|---|---|
| Brixey et al. (2009) | Dual channel volumetric general infusion pump | H, I |
| Carayon, Hundt, and Wetterneck (2010a) | Smart pump (one brand) | H, I, M |
| Carayon et al., (2007) | Bar code medication administration (BCMA) | H, I, M |
| Nemeth et al. (2009) | General infusion pumps (four different unnamed brands) | H, I, M |
| Chan et al., (2012) | -Synergy ® linear accelerator system (Elekta Medical)<br>-Desktop Pro™ 7 control system<br>-XVI™ cone beam imaging system<br>-iViewGT™ megavoltage system<br>-MOSAIQ™ record and verify system | H, I, M |
| Chiu, Vicente, Buffo-Sequeira, Hamilton, & McCrindle (2004) | Pacemaker programmer interfaces (six different unnamed brands) | I only |
| Etchells et al. (2006) | General infusion pump interfaces (two different unnamed brands) | I only |
| Fairbanks, Bishop, Marks, & Shah (2007) | Lifepak 10 and Lifepak 12 manual cardio-defibrillator devices | H, I, M |
| Garmer, Liljegren, Osvalder, & Dahlman (2002) | Infusion pump interface comparison of existing and new prototype | I only |
| Ginsburg (2005) | General infusion pump interfaces (two different unnamed brands) | I only |
| Graham et al. (2004) | One three-channel infusion pump (unnamed brand) | I only |
| Lin et al., (1998) | Graphical simulation of Abbott Lifecare 4100 PCA plus II infuser interface and prototype | I only |
| Lin, Vicente, & Doyle (2001) | Graphical simulation of Abbott Lifecare 4100 PCA plus II infuser interface and prototype | H, I, M |
| Liu, Tech, & Osvalder, (2004) | Numerical ventilator display and graphical user interface (GUI) prototype | I only |
| Nemeth et al. (2009) | General infusion pumps | H, I, M |
| Nunnally & Bitan (2006) | Infusion pumps (four different pumps from four different unnamed brands) | H, I, M |
| Obradovich & Woods (1996) | Infusion pump (unnamed brand) | H, I, M |
| Rogers, Mykitshyn, Campbell, & Fisk (2001) | Blood glucose meter | H, I, M |
| Trbovich, Pinkney, Cafazzo, & Easty (2010) | Infusion pump (three pumps from two unnamed brand) | H, I, M |
| Turley, Johnson, Smith, Zhang, & Brixey (2006) | Infusion pump operation manuals (five volumetric pumps from three unnamed brand) | I only |
| Zhang, Johnson, Patel, Paige, & Kubose (2003) | Infusion pump (two 1-channel pumps from two unnamed brand) | I only |

***Methods to identify problems.*** Heuristic evaluation provided the best option to investigate device adherence to established principles. Seven heuristic evaluations systematically compared the medical devices to established usability design principles or heuristics (Tables 3 and 4), including three studies that conducted heuristic evaluations using a single method approach (Graham et al., 2004; Turley et al., 2006; Zhang et al., 2003), and three studies that conducted heuristic evaluations in conjunction with other usability methods (Chan et al., 2012; Chiu et al., 2004; Etchells et al., 2006; Ginsburg, 2005). Observations and surveys were used to identify usability problems. These sources of data helped to explain the source of usability problems (Carayon, Wetterneck, et al., 2007; Chan et al., 2012; Obradovich & Woods, 1996), and user perceptions of usability problems (Caryon et al. 2010; Carayon, Wetterneck, et al., 2007; Chiu et al., 2004; Etchells et al., 2006; Obradovich & Woods, 1996; Rogers et al., 2001). The use of multiple methods to identify usability problems in order to triangulate data is supported in the literature (Garmer, Liljegren, Osvalder & Dahlman, 2002; Thyvalikakath, Monaco, Thambuganipalle, & Schleyer, 2009). Studies show that design characteristics impacted users and potentially patient safety (Hvannberg, Law & Lárusdóttir, 2007; Jaspers, 2009).

Table 3. Comparison of Nielsen (1993) and Shneiderman (1992; 1993) heuristic sets.

| **Nielsen (1993, p. 19) usability heuristic set.** | **Shneiderman (1992; 1998) eight golden rules** |
|---|---|
| • Simple and natural dialog<br>• Speak the users' language<br>• Minimize user memory load<br>• Consistency<br>• Clearly marked exits<br>• Shortcuts<br>• Good error messages<br>• Prevent errors<br>• Feedback<br>• Help and documentation | • Offer information feedback<br>• Support internal locas of control<br>• Reduce short-term memory load<br>• Strive for consistency<br>• Design dialogs to yield closure<br>• Enable frequent users to use shortcuts<br>• Permit easy reversal of actions<br>• Offer error preventions and simple error handling |

Table 4. Heuristic principles and definitions (Nielsen & Mack, 1994).

| **Heuristic (*variable*)** | **Definition** |
|---|---|
| Error prevention *****(*error*) | Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action. |
| Consistency and standards *****(*consistency*) | Users should not have to wonder whether different words, situations, or actions mean the same thing. Standards and platform conventions should be followed. |
| Recognition rather than recall *****(*memory/recognition*) | Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate. |
| User control and freedom (*control*) | Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo. |
| Match between system and the real world (*match*) | The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. The device should follow real-world conventions, making information appear in a natural and logical order; match the model the users have about the system. |
| Visibility of system status (*visibility*) | The system should always keep users informed about what is going on, through appropriate feedback or display within reasonable time. |
| Flexibility and efficiency of use (*flexibility*) | Accelerators -- unseen by the novice user -- may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions. |
| Aesthetic and minimalist design (*aesthetic*) | Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility. Extraneous information is a distraction. |
| Help users recognize, diagnose, and recover from errors (*recovery*) | Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution. Messages should allow all users to understand the nature of the error, learn, and recover from errors. |
| Help and documentation (*help*) | Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large. Help should be context sensitive. |

* Indicates heuristic principle most frequently violated in studies reviewed.

***Impact of usability problems.*** In the studies reviewed, the heuristic principles most frequently violated in studies reviewed were error, consistency, and memory (Chan et al., 2012; Chiu, Vicente, Buffo-Sequeira, Hamilton & McCrindle, 2004; Etchells et al., 2006; Graham et al., 2004; Turley, Johnson, Smith, Zhang & Brixey, 2006; Zhang, Johnson, Patel, Paige & Kubose, 2003). From Table 3, violations in consistency and memory affected user subjective workload and led to user confusion and errors. Also, violations of heuristic principles required short-term interventional strategies, such as user training to mitigate error or to guide hospital-based procurement decisions, when usability problems were severe; other problems required redesign, modification, or FDA recall (FDA, 2000).

**Environment.** Characteristics of medical device environment, specifically the physical and social environment, impacted usability in the studies reviewed. Sources of stimuli in the environment created visual or auditory distraction and/or interruption that competed with users' ability to effectively and efficiently interact with medical devices. Common sources of stimuli found in the literature were: ambient lighting, general noise level, clutter, and alarms (Brixey, Zhang, Johnson & Turley, 2009; Carayon et al., 2007), potential noise from helicopters or ambulances (Fairbanks, Caplan, Bishop, Marks & Shah, 2007), and inter-personal communication (Carayon et al., 2007; Ginsburg, 2005). While none of the studies measured the effects of these stimuli during user interactions with medical devices, nonetheless, they were important in describing the interactions and distractions during medical device use (Li, Magrabi, & Coiera, 2012; Rivera-Rodriguez & Karsh, 2010). After exhaustive searching was completed, there were no studies that described or measured the impact of these stimuli on user-device interaction.

This section reviewed the literature relating to the left side of the SMCIMD (Figure 5), system performance input factors, which were the interaction of nurse-user, device, task, and environment components as a system. Though numerous studies explored the impact of individual components on the left side of the study model (labeled nurse, device, task, and environment), no single study focused on the impact of internal and external stimuli between users and devices; specifically, the relationship between the nurse and medical devices, which is impacted by the force of factors within a demanding work environment.

**Nurse Interaction Processes**

The second section of this review is related to the center of the SMCIMD, which focuses on nurse interaction with medical devices and nurse interaction processes. These components of the system were the human information processing and attentional resources. Therefore, this section will review the literature relevant to human performance that resulted in the system performance outposts. This section will review: interruption in nursing practice, interruption frequency during medication administration, the nurse-PCA interaction, and the outcomes of the nurse-PCA interaction as it related to interruption.

**Interruption in nursing practice environment.** Interruption and distraction were two concepts that have been used interchangeably to describe an agent or event that shifts an individual's attention (Biron et al., 2009). An interruption was the human experience that created discontinuity in task performance situated within a specific context (Brixey et al., 2007). Distracters have been differentiated as precursors of interruption and as stimuli irrelevant to an individual's primary task that resulted in a break in attention or primary task activity (Biron et al., 2009; Healey, Sevdalis & Vincent, 2006). Sources of distraction could be internal, such as

the voluntary direction of a nurse's attention to another or a dual task, or external, as in an environmental auditory or visual distractions from phones, pagers, and equipment. Nurses may have ignored internal and external distractions; however, no current nursing or healthcare literature has been identified that described the characteristic intensity of external auditory or visual distractions or the subsequent impact of the interruption intensity. Once perceived, distractions resulted in interruption of a primary task with resumption later, interruption of a primary task with failure to resume the primary task later, or dual tasking of the primary task and a subsequent secondary task (Rivera-Rodriguez & Karsh, 2010).

**Interruption frequency during medication administration.** Interruptions most frequently occurred when nurses were preparing medications (Biron et al., 2009; Hall et al., 2010). Moreover, interruptions that occurred during medication administration were particularly dangerous. Interruptions significantly increased the risk of medication administration errors (Carlton & Blegen, 2006; Westbrook, Woods, et al., 2010), and when such errors occurred with interruption, they were more frequently characterized as serious, potentially leading to patient injury (Westbrook, Woods, et al., 2010). Interruptions may have been continuous or intermittent. Also, interruptions may have been of low, moderate or high intensity.

Interruptions during medication administration were most commonly measured using direct observation (Biron et al., 2009; Li, Magrabi, & Coiera, 2012; Rivera-Rodriguez & Karsh, 2010; Westbrook, Coiera, et al., 2010). Based on 14 observational studies, Biron et al. (2009) analyzed interruption frequency during medication administration; interruptions ranged from 0.8 to 41.8 events per hour (median 6.4). Westbrook et al. (2010) similarly reported that 53.1% of observed medication administrations ($n$ = 4271) were interrupted. Also, interruption frequency was significantly associated with task (medication administration) failures and errors. Logistic

regression showed the effect of interruption doubled when comparing zero to four interruptions in a single administration (Westbrook, Woods, et al., 2010). In addition to the risk of error, interruptions negatively affected nurses' performance, resulting in decreased task completion, increased errors, and increased perceived subjective workload (Redding & Robinson, 2009; Rivera-Rodriguez & Karsh, 2010).

Nurses described interruption and distraction as contributing factors to medication administration errors in several studies (Dickinson, McCall, Twomey, & James 2010; Jennings, Sandelowski & Mark, 2011; Potter et al., 2005). For example, interruption in the small, confined medication preparation areas was unavoidable and led potentially to errors (Dickinson et al., 2010). Searching for and implementing numerous devices contributed to an interruption-based environment and affected the medication administration processes (Jennings et al., 2011). In both studies, multi-tasking during medication administration processes was necessary to meet the temporal demands of the task. Device alarms and communication devices created an environment where nurses needed to be constantly available, creating an environment where nurses became unable to avoid distraction and interruption during critical safety processes of medication administration.

***The nurse-PCA interaction: Interruption and medical device use.*** Although medical devices are frequently used by nurses during medication administration, few studies reported the effects of interruptions on nurses' complex device use or interactions (Westbrook, Woods, et al., 2010). Nurses were frequently interrupted during medication administration, while using systems and devices such as automated dispensing machines, infusion pumps, and patient-controlled analgesia systems (Carayon et al., 2007; Ginsburg, 2005; Lin et al., 1998). Using quantitative approaches, these studies showed the negative effects of usability problems on the nurses'

accuracy, time to complete tasks, satisfaction, and subjective workload when interacting with the devices. These studies also observed distractions and interruption during device use in a naturalistic setting, but the studies did not measure the type or frequency of distracters and interruptions observed, or the effects of interruption on the nurses' interaction with the device. Therefore, nurses' perceptions of the impact of interruptions or intensity of interruptions were not considered in these studies. These studies suggested that future studies should consider the combined and cumulative effects of usability and interruption on nurses within their work environment.

Nurses have also reported that medical devices were overly complex and contributed to errors (Jennings et al., 2011; Treiber & Jones, 2010; Zuzelo, Gettis, Hansell & Thomas, 2008). Nurses' perceptions were congruent with studies that reported the complexity of devices negatively impacted health professionals' subjective workload and ultimately patient safety (Patel & Currie, 2005; Patel & Kaufman, 1998; Potter et al., 2005).

This section of the review was related to the center of the SMCIMD, which focused on nurse interaction processes, which were the components of a system loop that related to the aspects of human information processing and attentional resources. The studies concentrated on the changes within nurses' work environments, which have negatively impacted patient care. Specifically, the aforementioned research explored the negative aspects of frequent interruption and nurses' absolute need to multitask within their environment.

**System Performance Outputs**

The third section of this review related to the right side of the SMCIMD (Figure 5), which examines the factors of system performance outputs. System performance outputs were

defined as the results of the system performance input and nurse interaction processes as components within a system. This section will review the literature relevant to system performance outposts in terms of usability measures as system outcomes through effectiveness, efficiency, user satisfaction, and subjective workload.

**Usability measures.** Of the studies reviewed, ten studies quantified measures of usability. Appendix A, which can be found at the end of the review, was a literature synthesis table that describes key findings, and Table 5 briefly summarizes these measures according to effectiveness, efficiency, and satisfaction. Four studies reported performance measures and user satisfaction when comparing two or more medical devices (Garmer et al., 2002; Lin et al., 1998; Lin, Vicente, & Doyle, 2001; Liu, Tech, & Osvalder, 2004). Furthermore, data showed that new interface devices better supported user needs and limitations.

**Effectiveness.** Multiple studies compared existing device design to an improvement in new design with a higher accuracy rate (Garmer et al., 2002; Lin et al., 1998; Lin et al., 2001). Higher accuracy also aided in procurement decision making when comparing three pump choices (Ginsburg, 2005). Two studies reported measures of effectiveness using accuracy and task completion (Fairbanks et al., 2007; Trbovich et al., 2010). Limits in accuracy and completion demonstrated risks to patient safety, and the failure of a device to meet user needs and prevent errors through adherence to design principles, which also created risks to patient safety (Hornbæk, 2006).

Table 5. Summary of variable measured.

| Source | Effectiveness | | Efficiency | | | Satisfaction |
|---|---|---|---|---|---|---|
| | Task Completion | Accuracy | Time | Mental Effort | Interface patterns | Perception of Satisfaction |
| Carayon, Hundt, & Wetterneck (2010) | | | | | | ✓ |
| Chiu, Vicente, Buffo-Sequeira, Hamilton, & McCrindle (2004) | | | | | | ✓ |
| Fairbanks, Bishop, Marks, & Shah (2007) | ✓ | ✓ | | | | ✓* |
| Garmer, Liljegren, Osvalder, & Dahlman (2002) | | ✓* | ✓ | | | ✓ |
| Ginsburg (2005) | | ✓ | | | | ✓ |
| Lin et al., (1998) | | ✓ | ✓ | ✓ | | ✓ |
| Lin, Vicente, & Doyle (2001) | | ✓ | ✓ | ✓ | | ✓ |
| Liu, Tech, & Osvalder (2004) | | ✓ | | ✓ | | ✓ |
| Nunnally & Bitan (2006) | ✓ | | | | ✓ | |
| Trbovich, Pinkney, Cafazzo, & Easty (2010) | ✓ | ✓ | | | | |

*Indicates user-centered approaches.

**Efficiency.** A new PCA pump design improved task completion time significantly from 260 seconds to 188 seconds ($p<.05$), by implementing a user-centered design (Garmer et al., 2002). Other studies reported both an improvement in time to complete tasks and less mental effort exerted with an improvement in device design (Lin et al., 1998; Lin et al., 2001). Mental effort data was collected after each experiment using the NASA-TLX, an established multi-

dimensional assessment (Hart & Staveland, 1988). Task completion time and subjective workload improved significantly with the new device designs.

**User Satisfaction.** In studies comparing two or more devices, users preferred the device that improved ease of use, met user expectations (Carayon et al., 2010b; Fairbanks et al., 2007; Ginsburg, 2005; Lin et al., 1998; Lin et al., 2001; Liu et al., 2004), and devices designed using a user-centered approach (Fairbanks et al., 2007; Garmer et al., 2002). Studies reviewed used non-standardized measures despite the availability of reliable, valid measures (Hornbæk, 2006).

***The impact of interruption on the human-device (PCA) interaction.*** Using three established measures of usability, the HF/E evidence described the effects of interruption on four components of human-device interaction, as described by the International Organization for Standardization (ISO) 9241-11 standard. The extent to which a product could be used by specified users, such as nurses, helped to achieve specified goals with efficiency, effectiveness, and satisfaction (ISO, 1998). Usability measures were the standard for measuring outcomes during the human-device interaction with medical devices (U. S. Food and Drug Administration, 2011; FDA, 2000). Despite the standards and past successes in other high-risk, high consequence industries, few HF/E studies have involved nurses interacting with medical devices; further, nursing research has not quantified the effects of interruption or employed established usability measures.

*Measure of Efficiency with PCA interaction.* Measures of efficiency were temporal and human performance related. Given the time-dependent nature of nursing processes, such as medication administration, the time it took to complete tasks was an important aspect of efficiency. Task completion time was the most common temporal measure (Hornbæk, 2006). Mean task completion times for interrupted medication administration tasks were significantly

longer than non-interrupted tasks (Trbovich, Prakash, Stewart, Trip, & Savage, 2010). Interruption may have resulted in faster task completion times, but the rate of speed occurred at the expense of increased perceived subjective workload (Li et al., 2012). These divergent findings supported the need for additional study.

*Measure of subjective workload.* Subjective workload, the human mental or cognitive effort expended during human-device interaction, has emerged from the interaction between the requirements of a task, the circumstances under which tasks were performed, and the skills, behaviors, and perceptions of the human user (Hart & Staveland, 1988). Only one nursing study, Kataoka, Sasaki, & Kanda (2011) addressed the impact of interruption on nurses' subjective workload during infusion pump use. The study found that nurses experienced an increased subjective workload due to time pressure during shortened infusion pump operation. Other studies found subjective workload to be negatively impacted by interruption (Palese et al., 2009; Redding & Robinson, 2009) and poor medical device usability (Lin et al., 1998; Lin, Vicente & Doyle, 2001). According to these studies, the combined or cumulative effects of poor device usability and interruptions have contributed to errors during nurses' interaction with PCAs.

*Measure of effectiveness.* The accuracy and completeness of the human-device interaction with a medical device determined effectiveness (ISO, 1998). Accurate completion of programming tasks was critical to achieving any patient safety goal. No nursing studies have reported accuracy of PCA programming task completion, especially related to interruption. However, it was suggested from simple computer-based tasks that interruptions of only 2.8 seconds could double the error rate, and interruptions of 4.4 seconds could triple the rate of error (Altmann, Trafton, & Hambrick, 2013). These findings supported concerns regarding the impact of interruption on nurses' PCA programming tasks.

*Measure of satisfaction.* Measures of satisfaction described how users feel about a system or device, which linked aspects of the human-device interaction, such as interface design, tasks, or environment that did not meet the user needs, limitations, or expectations (Fairbanks, Caplan, Bishop, Marks, & Shah, 2007; Garmer, Liljegren, Osvalder, & Dahlman, 2002; Lin et al., 1998). User satisfaction helped to identify problems with displays, controls, and operation that were frustrating, stressful, or overwhelming, and may have limited safe use (Bennette, Dawoud & Maben, 2010; Chiu, Vicente, Buffo-Sequeira, Hamilton & McCrindle, 2004; Palmer et al., 2013). Nurses' satisfaction with medical devices and other technology described the positive and negative impact on nursing practice, such as the increased risk of error as a result of overly complex design and programming tasks (Marini, Hasman, Huijer, & Dimassi, 2010). Perceptions of user satisfaction provided valuable insight into the quality of an interaction and how users were impacted by device interactions. Standardized measures of satisfaction were recommended for improved reliability and validity (Hornbæk, 2006).

The success of the conceptual framework was measured by efficiency, effectiveness, satisfaction, and subjective workload within the system performance output. Previous research found that system performance output suffered due to interruptions within the work environment, including the limitations of certain device usability. Findings from this review delineated a gap in current research with regard to nurses' perspectives of the impact of interruption frequency in relation to the quality of task completion and increase in subjective workload.

## Current Medical Device Usability

Conceptually and practically, users were central to usability. Despite the user's central role, usability was realized through system-level analysis and adherence to established medical

device design principles. Human factors and system-level analyses effectively identified, described, and explained medical device safety concerns that may have led to errors, injuries, and recalls (Bagian, 2012). Current medical device usability studies demonstrated that existing devices may not have been designed with a full consideration of the complex healthcare systems where devices are used (Carayon et al., 2007; Chan et al., 2012; Fairbanks et al., 2007). They may have not been designed according to established design principles (Graham et al., 2004; Turley et al., 2006; Zhang et al., 2003). Medical device design flaws may not be identified until after FDA approval or sale. A thorough review of the literature provided evidence that once devices were redesigned, new designs improve efficiency, effectiveness, subjective workload, and satisfaction (Graham et al., 2004; Lin et al., 1998; Lin et al., 2001; Liu et al., 2004). The studies specifically highlighted the importance of user-centered design and inclusion of a system-level perspective, which led to improved medical device use and patient safety. However, research addressing medical device usability with attention to the system of use should be pursued until medical devices no longer contribute to patient safety issues.

Issues and difficulty in the measurement of usability of medical devices were evident in these studies. Multiple definitions of usability exist and each definition was comprised of various attributes, dimensions, or component of usability (Folmer & Bosch, 2004). Contributing to measurement problems, usability could only be measured indirectly (Hornbæk, 2006), necessitating reliable, valid measure of development and use. There was a debate regarding the need to measure general usability, implementing measures such as the System Usability Scale (SUS) (Bangor, Kortum, & Miller, 2008), or context specific usability (Liljegren, 2006; Trivedi & Akheela Khanum, 2012). Multiple approaches described usability problems and measure usability; each approach had advantages and disadvantages (Jaspers, 2009).

**Limitations of This Review**

Studies reviewed were limited to those that were published and peer-reviewed literature. The search-screening strategies (Figure 6) may have limited sources. Other sources may be available in other electronic databases and non-peer reviewed work. The integrated review method (Whittemore & Knafl, 2005) was selected because the sources were from multiple disciplines, employing a variety of methods, devices, and measures. The recommended strategies were employed to improve rigor, however, the synthesis of these diverse sources may benefit from the insight of a second reviewer.

**Limitation of Current Studies**

For studies reviewed (Appendix A), limitations were identified in sample size, lack of conceptual/theoretical framework, and limitations in author reports of reliability and validity. Transferability or generalizability relating to sampling was a limitation in most studies. Studies lacked details regarding usability approaches and many did not acknowledge methodological and design limitations, reliability, and validity. Despite these limitations, this review of medical device usability analysis findings could guide future research.

# CHAPTER THREE: METHODOLOGY

The purpose of this study was twofold: to quantify the impact of interruption frequency on registered nurses' performance, satisfaction, and subjective workload during PCA interaction and determine nurses' perceptions of the impact of interruption frequency. Included are sample inclusion criteria, participant recruitment strategies, protection of human subjects, the study setting and details about the researcher. Additionally, this chapter describes the study setting and provides details about the researcher.

## Research Method and Design Appropriateness

### Design

A mixed-methods approach was used for this research. First, an experimental repeated measure design was used to quantify the impact of interruption frequency for aims one and two of the study. After each experiment, semi-structured interviews were used to collect data that was analyzed to determine nurses' perceptions of the impact of interruption frequency and interruption intensity on their PCA interactions for aim three of the study.

In this study, there was one independent variable (interruption frequency) with three condition levels. Participants were randomized to each condition. There was one counterbalanced between-subject factor, task order with six possible groups (see Table 6). The study measured four dependent variables: efficiency (task completion time), effectiveness (accuracy), subjective workload (NASA-TLX), and satisfaction.

The experimental repeated measured design is best when participants such as experienced nurses will be difficult to recruit and when tasks are complex, as with working with advanced medical devices. The design is supported when multiple, different treatments are utilized,

requiring a smaller sample size and ensuring all participants receive treatment in all conditions, effectively isolating individual differences that may occur within the sample (Lazar, Feng, & Hochheiser, 2010).

Table 6. Between-subject group task order variations.

| | |
|---|---|
| 1-2-3 | 1-3-2 |
| 2-1-3 | 2-3-1 |
| 3-1-2 | 3-2-1 |

Semi-structured interviews provided access to nurses' perceptions as a lens through which the researcher obtained unobservable data from high-risk situations that may explain the impact of interruption on nurses' PCA interactions. Qualitative descriptions using nurses' own words are the best method to determine the impact of interruption frequency and interruption intensity given the current knowledge limitation. This study followed a constructivist paradigm which supports building knowledge from multiple data sources to improve understanding of a problem (Lincoln & Guba, 1985; Matney, Brewster, Sward, Cloyes, & Staggers, 2011). The principal investigator's (PI) experience as a registered nurse and knowledge areas were considered an advantage aiding in data collection and data analysis. However, the PI's experience with qualitative methods was a limitation which required reliance on the dissertation committee or other experts. The use of quantitative and qualitative data provided a better understanding of the impact of interruption frequency on nurses' PCA interactions than either research approach alone.

**Definitions**

The tables below provide the definitions of terms used in this study.

Table 7. Theoretical and operational definitions.

| Study variables | Conceptual Definition | Operational Definition/Measurement |
|---|---|---|
| (1) Efficiency | Time and human resources consumed in order to complete tasks. | Task completion time measured in seconds to complete each PCA programming task. Total task completion time is the total of all task times for each participant. |
| (2) Subjective workload | The human mental or cognitive effort expended during human-device interactions. | Subjective workload as measured using the total score on the NASA-TLX using six subscales of mental demand, physical demand, temporal demand, performance, effort level, and frustration level. |
| (3) Effectiveness | The level of accuracy at which users achieve specified tasks. | Accuracy as measured categorically accurate or inaccurate when users complete each PCA task. |
| (4) Satisfaction | Attitudes toward the use of a system. | Measured with one item (frustration) on the NASA-TLX and one item (item 2) on the semi-structure interview. |

Table 8. Demographic and control variables.

| Demographic | Variables |
|---|---|
| – Age | – Participant's age |
| – Gender | – Participant's gender |
| – Ethnicity | – Participant's ethnicity |
| – Vision | – Participant's vision status |
| – Colorblindness | – Participant's colorblindness status |
| – Basic nursing education | – Point of entry into basic nursing practice; first nursing degree achieved for practice as a registered nurse |
| – Educational achievement | – Participant's highest degree earned to date |
| – Nursing Experience | – Number of years as a practicing registered nurse |
| – Work status | – Number of hours worked per week on average (part or full time) |
| – Employer hospital size | – Number of beds at hospital of employment |
| – Years at employer hospital | – Number of years participant has worked at the hospital of employment |

| Demographic | Variables |
|---|---|
| – Computer expertise | – Self-rated level of comfort using a computer |
| – Internet use | – Self-rated frequency of internet use |
| – PCA expertise | – Self-rated level of comfort using a PCA |
| – PCA use | – Self-rated frequency of PCA use |
| – Baxter PCA II Pump Experience | – Participant's experience with the pump used in this study |
| – Unit patient mix | – Unit description of patient care mix |
| – Year experience on unit of employment | – Number of years participant has worked on the unity at the hospital of employment |
| – Certification status | – Determination of specialty certification status |
| **Control Variables** | |
| – Interruption frequency | – The rate of auditory or visual stimuli perceived by a nurse operationalized in condition A, condition, B, and condition C. |
| – Work setting | – Subjects limited to medical-surgical RNs via sampling |
| – Work qualifications | – Subjects limited RNs who have at least 6 months in adult medical-surgical setting via sampling |
| – Work status | – Subjects limited to RNs working at least 24 hours per week via sampling |

**Research Questions and Hypotheses**

This specific aims of this study are listed below. The central hypothesis of this study was that interruption frequency during nurses' patient-controlled analgesia device interaction will affect nurses' performance efficiency and effectiveness, subjective satisfaction, and perceived subjective workload.

*Aim #1. Determine the impact of interruption frequency on nurses' PCA performance.*

- *Research question #1:* What is the effect of interruption frequency on the efficiency and effectiveness of medical-surgical nurses' PCA use?
- *Hypothesis #1:* Increased frequency of interruption will have a negative effect on nurses' performance efficiency (EF1-task completion time) and effectiveness (A1-accuracy).

*Aim #2: Determine the impact of interruption frequency after PCA interactions on medical-surgical nurses' subjective satisfaction and subjective workload.*

- *Research question #2:* What is the effect of interruption frequency on medical-surgical nurses' subjective satisfaction and subjective workload with PCA use?
- *Hypothesis #2.* Increased interruption frequency will decrease nurses' subjective satisfaction and increase subjective workload.

*Aim #3: Determine nurses' perceptions of the impact of interruption frequency on nurses' PCA interactions.*

- *Research question #3*: What are medical-surgical nurses' perceptions of the impact of interruption frequency during PCA interactions?

**Population**

A purposive sample of nine licensed registered nurses (RN) experienced in adult medical-surgical acute care were recruited. The following RNs were eligible for inclusion: (a) employed for 24 or more hours per week on average in a medical-surgical unit; (b) had at least six months experience in adult medical-surgical nursing; and (c) indicated (self-report) PCA use at least four shifts per month. The following RNs will be excluded: (a) RNs who work less than 24 hours per week on average; (b) RNs with less than six months of experience as an RN; (c) RNs who do not use a PCA at least four times per month; and (d) RNs whose primary unit of employment is high acuity areas such as intensive and critical care, or specialties areas other than medical-surgical.

**Limitations of the sampling method.** The volunteer nature of the sample may bias results toward RNs who have specific perceptions and want to verbalize perceptions regarding

interruptions and/or PCA use. Every effort was made to recruit RNs who appeared comfortable interacting and expressing their full experience to achieve the aim of the study.

**Power analysis and sample size.** Usability testing involves a specific medical device with a heterogeneous sample of users (Wicklund, Kendler, & Strochlic, 2011). Reliability of results is linked to variability in sample characteristics and sample size when studying human-device interactions (Thyvalikakath, Monaco, Thambuganipalle, & Schleyer, 2009). Sample size in comparable empirical usability studies ranges from 6 to 24 (Liu, Tech, & Osvalder, 2004; Trbovich, Pinkney, Cafazzo, & Easty, 2010). Power of .80 is acceptable in usability testing and effect size between small and moderate are often not practically meaningful in usability testing (Nielsen, 1997; Salvendy, 1997). Known barriers to medical device usability testing are the difficulty of recruiting qualified medical device users (Wicklund et al., 2011) and missing data can be an issue during the data analysis. Small sample size is a limitation in many current studies (Campoe, 2013a).

Assuming a power of .80, alpha of .05, and within-subjects correlation of .90, a sample size of 7 is needed to detect a moderate effect size (Cohen's $f$=.25) when ANOVA-RM, within factors is used for *a priori* calculated with three continuous dependent (efficiency/task completion time; subjective workload; and satisfaction) variables using G*Power version 3.1.5 (Buchner et al., 1997). Actual power of .82 is estimated. Anticipating an estimated 10% attrition and 10% missing data as a cutoff, a total sample size of 9 nurses was sought to balance feasibility, current shortcomings in comparable studies, and minimum sample size needed to detect moderate effect. An expert statistician confirmed the results of the power analysis prior to the start of the study.

**Sampling Design and Participant Selection**

A recruitment flyer and emails were distributed to medical surgical hospital nurse managers and leaders in west central Florida requesting that they assist with recruitment of qualified participants. Flyers were emailed to the pool of available nurses and posted in central areas for each medical-surgical unit and hospital. Interested nurses contacted the PI. Nurses who meet inclusion criteria were accepted and scheduled for the experiment. Prior to commencement of the study, participants were randomly assigned to condition using a computer-generated randomization. Task order was counterbalanced to each condition, whereby each participant will ultimately complete PCA programming task in each condition. Final number of participants recruited was 9.

**Informed consent.** The consent process was initiated only after IRB approval at both the University of central Florida and Nova Southeastern University. Registered nurses who agreed to participate were scheduled to complete the informed consent process at Nova Southeastern University on the day of the scheduled experiment. The PI completed consent for all participants in person. The consent process was conducted in English in a private meeting room at Nova Southeastern University. The consent process was verbal informing participants as well as in writing of the study purpose, expectations of participants, as well as study risks and benefits. Informed consent process took approximately 15 minutes per participant. The consent (Appendix B) specifically described the following:

- Participation will be audio and video recorded and later transcribed resulting in a paper and electronic versions of the nurses' participation.
- Interviews will be audio recorded then alter transcribed resulting in a paper and electronic versions of the nurses' interview.

- The data collected will be used solely for the purposes of the study by the PI and supervising faculty.
- Participants will be offered the option allow the PI to present non-identifiable audio or video data at research conference.
- The demographic questionnaire will be completed only after the informed consent is completed.

**Confidentiality.** A specific process and special provisions were taken to maintain confidentiality. The study collected self-report and human performance data from registered nurses including audio-video recordings. Data was de-identified to provide anonymity, privacy, and confidentiality of participants. Self-reported data and audio-video recorded data were assigned a unique identifier to disassociate identify from data. The PI was responsible for the overall planning and implementation of this study ensuring the scientific integrity of all aspects of the project and data. Data were stored and locked at the Nova Southeastern University College of Nursing in the PIs' office. Only the PI had immediate access to original and stored data. Consent forms and collected data were stored in separate locked locations at the PI's office.

**Setting.** The study was conducted in the Nova Southeastern University Anesthesia Assistant simulation laboratory. Permission to use the simulation laboratory, medical equipment, devices, and technical support was secured in advance of the study. The simulation laboratory was arranged to simulate an in-patient medical-surgical nursing environment including beds, furniture, patient room phone, patient cell phone, television, general infusion pumps, PCA, pulse oximeter, and lighting. This simulation laboratory was viewable via one-way glass limiting intrusiveness of the PI. Existing audio and video-recording (AV) equipment was used.

**Procedures to protect privacy.** Participation in the study was voluntary. Multiple steps were taken to protect participant privacy. First, the physical location of Nova Southeastern University and the physical location of the high-fidelity simulation laboratory were not expected to pose a threat to participant privacy. Second, prior to the experiment, consent, collection of demographic data, and training were conducted in a private room, located next to the high-fidelity simulation laboratory. Only the PI had access to the participant during consent and training; no audio or video recordings were made. Third, during the experiment, data collection included audio and video recordings of the participants' performance and interaction with the PCA. Audio and video recordings were used to collect outcome data related to total task time and accuracy. Only the PI and members of the research team were present in person or via one-way viewable glass window. Only the PI and members of the research team had access to view the completed audio and video recordings. The semi-structured interview was audio-recorded. All verbal data from the experiment and interviews was transcribed into a text document. Video data was de-identified to provide anonymity and privacy to participants. Finally, demographic data, the study outcome measures, and audio-video recordings were necessary and fundamental to the study. Only data necessary to the planned research study was collected. Research reports aggregated data rather than report individual participant data.

## Measures and Instruments

Interruption frequency was the independent variable with three levels in this study. Four dependent variables were measured: efficiency (task completion time), effectiveness (accuracy), subjective workload (NASA-TLX), and satisfaction.

**Operationalization of the Independent Variable (IV) treatment conditions.**

Interruption frequency was defined as the rate of auditory or visual stimuli perceived by a nurse. In this study, interruption frequency had three pre-determined levels. All participants completed tasks in condition A first then each treatment condition was randomized. The conditions for interruption frequency are as follows: Condition A was free of interruptions. In most usability studies, participants complete tasks without distraction or interruption (Campoe, 2013a). Condition A represented the interruption-free testing environment as identifeid in current published medical usability analysis literature.

1. Condition B contained two planted interruptions per 10 minute task scenario. A systematic review of literature pooled data from 14 studies reporting interruption frequency and other characteristics of interruption (Biron, 2009). Interruptions ranged from 0.8 to 41.8 events per hour and the mean calculated interruption frequency was 6.7 to15 events per hour. Condition B represented the mean rate of agents or events that shifts nurses' attention based on current knowledge.
2. Condition C contained four planted interruptions per 10 minute task scenario. The effects of interruption doubled when comparing zero to four interruptions in a single task administration (Westbrook, Woods, et al., 2010). Condition C simulated the rate of agents or events found by Westbrook et al (2010) to double the risk of error.
3. Condition D contained six planted interruptions per 10 minute scenario. Biron (2009) found that interruptions ranged from 0.8 to 41.8 events per hour. Condition D represented the maximum range of 41.8 interruption events per hour that shifts nurses' attention identifeid in current literature.

Most empirical human factors usability studies are conducted in a simulated environment without realisms such as distraction, interruption, and other environmental factors that impact human performance (Fairbanks et al., 2007; Lin et al., 1998; Lin et al., 2001). Each condition simulated an experimental testing environment similar to published HF/E usability and related studies. In addition, participants completed each task without planned interruption at the end of training after orientation that will serve as a control or baseline for the experiment.

These high-fidelity experimental conditions were developed, using realistic scenarios that embed representative PCA programming tasks into the scenarios (Lazar et al., 2010; Maddox, Danello, Williams, & Fields, 2008; Rubin, 2008). The conditions and tasks were created based on PI expertise and current literature. Each IV condition was operationalized to simulate nurses' typical channels of interruption using a pre-recorded message to be played on the overhead intercom. The participant were cued to respond to and required to turn away from the patient and PCA to respond to computer questionnaire simulating interruptions. Computer simulated interruptions were used. Planting distractions and interruptions has been applied in similar usability studies (Carayon, 2010; Coursaris, Hassanein, Head, & Bontis, 2012; Magrabi, Li, Day, & Coiera, 2010; Prakash & Trbovich, 2012). Others studies have planted errors in infusion pump task conditions to measure error resolution (Trbovich, Pinkney, et al., 2010). The developed scenarios and tasks were validated by one nursing expert and one human factors expert. The scenarios were tested in the pilot with subsequent revision if needed. See appendix C for scenarios with programming tasks and interruptions.

**Dependent variables.** The FDA (2000) has recommended using reliable, valid measures during medical device usability testing. Established objective and subjective measures of

usability (Hornbæk, 2006) were selected for this study to achieve the aims of this study. Measures were collected during or after each condition.

*Efficiency* is an objective measure that was captured during simulated PCA programming. The task completion time (seconds) (EF1) was the time it takes to complete each PCA programming task. Total task completion time (seconds) (EF2) was the sum of task times, or the total time it took to complete all PCA programming tasks for each participant. Task completion time is a well-established measure of the interval between the time that participants first touch the PCA to initiate programming action (or sub-task) and the time of the last programming action, subtask or signal (Hornbæk, 2006).

*Effectiveness* is an objective measure that was captured during simulated PCA programming as accuracy. Accuracy (A1) is categorical variable whereby there is only one accurate outcome of each PCA programming task (Hornbæk, 2006). Completed task programming was categorized as error-free (no errors) or not error free for each tasks completed. An error log of all errors was recorded. The error log provided documentation of issues that users experienced resulting in tasks that are not error free and included a description of the error, a description of the task where the error occurred, the impact of the error if determined. The cause of the error was documented, if determined.

*Subjective workload* is a subjective measure of usability as well as measure of efficiency that was completed after programming tasks in each condition. This study licensed use of the NASA-Task Load Index (NASA-TLX) (Hart & Staveland, 1988). See Appendix D for the paper and pencil version of the NASA-TLX with subscale definitions and Appendix H for permission to use the NASA-TLX. This multi-dimensional assessment has six subscales measuring mental demand, physical demand, temporal demand, performance, effort level, and frustration level. The

six subscale can be further described in three dimensions: task- related, behavior-related, and subject-related (Hart & Staveland, 1988). The task- related factors describe the objective demands imposed by tasks on the user. Three subscales represent the task-related factors: mental demand, physical demand, and temporal demand. The subject-related factors describe the user's subjective response to task interactions. Two subscales represent the subject-related factors: effort and performance. Finally, the subject –related scale describes the psychological impact of the task demands, behavior, and performance on the user. The frustration subscale represents the subject-related factor.

Each subscale ranges from very low (0) to very high (10). See Appendix C for subscale definitions. The overall subjective workload score is a combination of the six dimensions (Hart, 2006). Items were summed with a higher score indicating higher perceived subjective workload. This study eliminated the pair-wise comparisons and used the unweighted NASA-TLX scores given that the procedures for weighting have limited benefit (Hendy, Hamilton, & Landry, 1993; Nygren, 1991).

The NASA-TLX has been used in the healthcare and for usability testing in the simulated environment using a variety of users, including nurses (Hart, 2006; Hoonakker et al., 2011; Weigl, Müller, Vincent, Angerer, & Sevdalis, 2012). NASA-TLX test-retest reliability of 0.77 has been reported (Hoonakker et al., 2011) and high concurrent validity (.73-.79). The NASA-TLX had high positive correlations (.97-.98) with comparable tools (Rubio, Díaz, Martín, & Puente, 2004). The tool is reliable and valid for subjective workload assessment in ICU nurses (Hoonakker et al., 2011).

*Satisfaction* was measured after completion of programming tasks using two measures. First, the NASA-TLX subscale of frustration level described how secure, gratified, content,

relaxed and complacent the user felt during the tasks versus insecure, discouraged, irritated, stressed and annoyed (Hart & Staveland, 1988). This measure was completed after each condition. Second, item two on the semi-structured questionnaire required the user to rate the impact of interruption frequency on satisfaction. The scale used a four point Likert-type scale ranging from no impact to high impact.

*Nurses' perceptions* of the impact of interruption frequency on their PCA interaction was a subjective measure captured after completion of tasks in all conditions using a semi-structured interview. An interview guide was developed by the principal investigator (PI) (see Appendix E). Questions one through three were developed to ascertain the participants overall perspective of interruption frequency, allowing the participant to rate their responses on a four point Likert-type scale ranging from no impact to high impact. After items one through three, open ended questions were used allowing the PI to guide the discussion and build rapport while the RN as questions then become focused exploring RNs' perceptions of the impact of interruptions during PCA use and participation in the experiment. Semi-structured interviews are a data collection method useful for triangulating data from observations and interviews (Martin, Norris, Murphy, & Crowe, 2008).

**Demographic data.** Demographic data was collected from participants. Nurse characteristics (e.g., age, years of nursing experience, nursing education level, frequency of PCA use, PCA programming experience, work-hours per week) will be reported descriptively and used to assess for differences and comparison in levels of education and experience.

**Device.** The patient-controlled analgesia (PCA) system for this study was the Baxter PCA II Pump (Model 2L3104), a syringe pump made for hospital use (Baxter Healthcare, 1993). The device (Figure 5) holds pre-filled or standard syringes and is programmed for medication

administration in milliliters or milligrams in one of different modes: PCA only, Basal plus PCA, and continuous basal. The keypad has three buttons plus numerals (0-9). The display holds 8 lines with up to 14 alphanumeric characters in a backlit LCD display. The device includes a locking syringe cover, lockable pole clamp, and keypad access codes to prevent theft, loss. The device requires tubing fitted with a specialty cassette for the device. The device is battery operated with a 9-volt alkaline battery or with AC power supply. The device dimensions are 13" H x 6.3" W x 2.8" D weighting 4.2 pounds. The Baxter PCA II Pump Operator's Manual details instructions for safe use and was used to develop PCA programming tasks.

**Data Collection Procedures**

**Participant training.** After consent process was completed, each participant received a brief 15-minute training session on the PCA pump including general functionality of the PCA and common programming tasks required in the experiment. Training was intended to give each participant, regardless of background or experience, a similar base for completing PCA programming tasks in the study, and was not intended to train the participant to the level expected of an expert user.

Participants entered the test environment and the facilitator provided an orientation to the setting, ensure settings on the PCA were accurate for the scenario, fit the participant with a microphone, then set and check the AV equipment. All participant interactions were audio and video recorded in their entirety, including full recordings of participant's performance of PCA programming tasks with screens, all interactions with the nearby computer screen and simulated patient, and activities occurring in the room.

Orientation included the explanation that work environment was to be interrupted as is usual in a nurse's work environment. The participant was instructed to complete PCA programming tasks throughout the experiment as if it were their own work environment with the goal of completing PCA programming efficiently and accurately.

Participants were instructed to acknowledge interruptions, and attend to each interruption. During each patient care scenario in the experiment, the participant was deliberately interrupted over the intercom. The participant heard a verbal interruption, "Excuse me, could you please assist me?" This message was intended to simulate the most common interruption in healthcare: interpersonal communication. This message prompted the participant to stop the PCA programming task and turn toward a computer screen that had been placed on a table approximately five feet from the participant. At the computer screen, the nurse responded to one survey question, submit a response, and then return to the PCA programming task.

Each interruption from the recording required the participant to stop the primary PCA task during programming to (a) acknowledge interruptions by turning away from the PCA and patient, (b) cognitively multi-task, then (c) attend to the interruption as warranted (Grundgeiger et al., 2010). After responding to the survey question, the participant was directed to return to the primary PCA programming task until completion.

At the conclusion of the training and orientation, each participant was presented with a verbal report and orders on the four patient scenarios (Appendix C) to be encountered during the experiment. The participant was permitted to take notes. The verbal report was intended to simulate the change of shift report that that commonly occurs prior to patient care or change of shift in the naturalistic setting. This verbal report was used to communicate current patient situation, background, assessment, and orders with recommendations.

**Experimental conditions and tasks.** After the verbal report, each participant was presented with the first scenario as condition A, which was free of interruptions. The participant was instructed to perform the PCA programming tasks per the orders. The participants were instructed to signal upon completion of each task. Upon completion of the tasks, the NASA-TLX was administered. These data functioned as a baseline data for the experiment.

Three treatment conditions (Conditions B, C, and D) and tasks (Appendix C) were randomized prior to the experiment. Once ready, participants performed PCA programming tasks in the first condition (B, C, or D) as randomly assigned. The participants signaled upon completion of each task. Once all tasks were completed in the second condition, the NASA-TLX was administered. Participants were required to take a five minute break to rest while the research assistant or PI prepared for the third randomly assigned condition and task scenario. Participants then completed the same procedures in the third and fourth randomly assigned conditions. A five minute required break separated the second and third condition. Immediately following completion of tasks in the fourth condition, a brief semi-structured interview was conducted with all participants. Total time to complete the experiment for each participant was anticipated to be 1.5 hours. Upon completion of study measures in four conditions and the interview, participants were given the incentive of a pre-paid $45 retail gift card. Participants were required to complete the experiment to receive the incentive. See Figure 7 for a diagram of the experimental procedures.

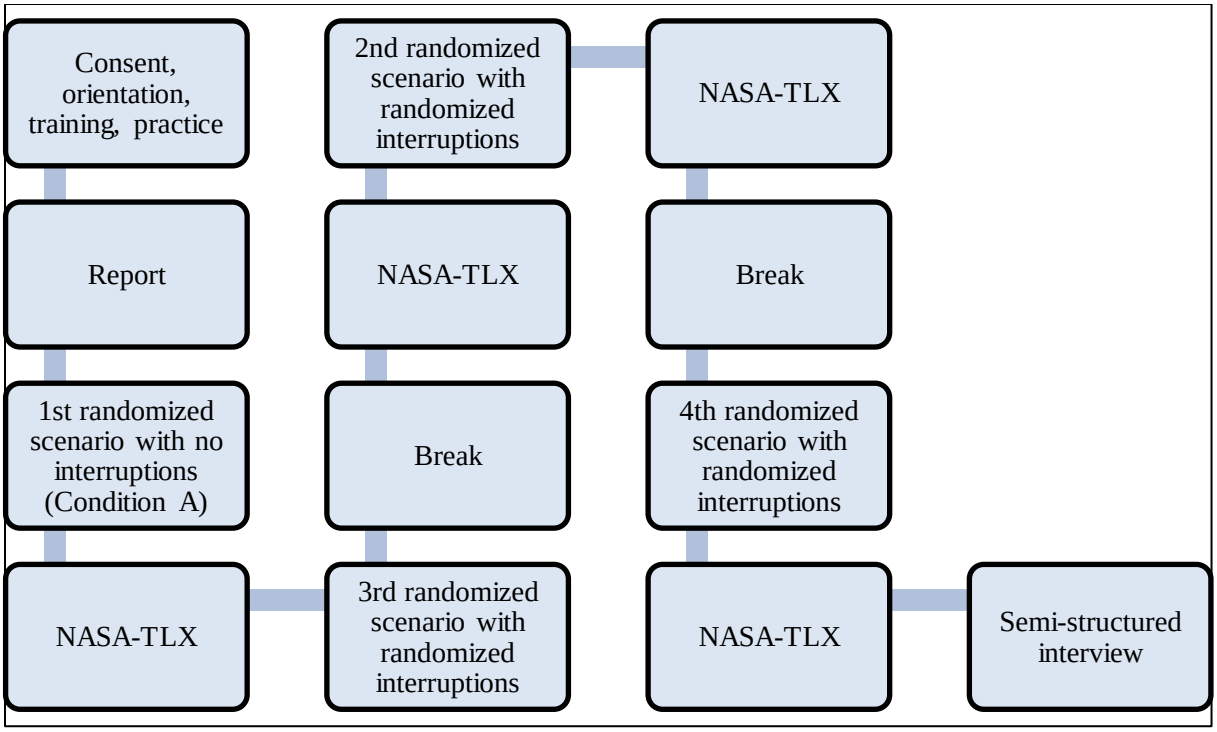



Figure 7. Diagram of the experimental procedures and interview

**Data collectors and training.** Data collection was conducted by the PI. She is experienced in medical device usability data collection (Campoe et al., 2012), and the high-fidelity simulation to be used as the study setting.

**Data collection.** The PI tested all study procedures and measures prior to the empirical study in a pilot study ($n$=3). There were three sources of data in this study: self-report instruments, nurse performance data, and semi-structured interview data. First, self-report data collected from the demographic questionnaire and NASA-TLX were entered into MS Excel spreadsheet, cleaned and prepared for export to the statistical software. Next, nurse performance data for efficiency (EF1- task completion time in seconds; EF2- total task completion time in seconds) and effectiveness (A1-accuracy) measures were collected during simulation using audio-video recordings. Audio-video data was automatically synchronized by the software and

time-stamped. Audio-video recordings were evaluated and data was abstracted for task completion time (EF1) in seconds by entering start and end time in seconds into a MS Excel spreadsheet. Task completion time (EF1) in seconds was calculated based use the audio-video time-stamp data as the start and stop time for the established measure. Total task completion time (EF2) is the sum total of all task times for each participant. Next, audio-video recordings were reviewed and coded for the measure of accuracy (A1) using the established measure for each task. Errors were described in the error log. Nurse performance data were entered into a MS Excel spreadsheet for data abstraction.

Finally, data from semi-structured interviews data were collected in a single interview after participants completed tasks in all conditions. The interview was digitally audio-taped and transcribed including notation of pauses, interruptions, and changes in voice, tone, and any noted emotion to ensure reliability of data. Transcripts were proofed for accuracy by the investigator. Transcript data was entered into Dedoose® (SocioCultural Research Consultants, 2013), a qualitative software program available for this study. A member of the dissertation committee experienced in qualitative data methods validated the data collection and entry into Dedoose®.

**Data management.** All data remained in a secure format and locked at the PI's office on campus at NSU to protect confidentiality. Regularly scheduled meetings were conducted by the PI with supervisory faculty via phone to promote communication and ongoing quality management once data collection began. Strategies to improve reliability of data collection and data analysis included the requirements of CITI training (PIs, collaborators, student), training on use the PCA device, and a pilot of data collection methods and procedures.

**Pilot.** The pilot was conducted prior to the empirical phase of this proposed study and was used to assess and refine the adequacy of all study procedures, training, and data collection

methods as well as evaluate quality of data yielded from study measures and semi-structured interviews. Nurses (*n*=3) from the described sample were recruited to participate in the pilot. During the pilot, the NASA-TLX computer version was not functioning consistently, so the paper-pencil version of the NASA-TLX was tested during the third trial of the pilot. As a result of the poor functioning of the NASA-TLX computer version, the paper pencil version was used for the study. No other substantive changes were made to the study as it was described.

**Threats to internal and external validity.** There were several potential challenges and limitations to this study. First, the within groups study design was limited because it is difficult to control for learning effects and there may be effects of fatigue after participating in multiple conditions. To account for learning effects, training was planned to be sufficient to allow time for participants to become familiar with the device and task. Participants were given a break in between the experimental conditions to limit effects of fatigue.

Participants were recruited from a specific region of west central Florida, convenient to the PIs and study site. Participants may not have be representative of other medical-surgical nurse populations. Finally, the high-fidelity simulation laboratory setting, nature and frequency of planted interruptions, and PCA programming tasks were potentially not representative of the setting in nurses. Subsequently participants may have committed more time to complete tasks or reacted differently to interruption than they would in their representative practice setting. The use of high-fidelity rather than low or no-fidelity is a trade-off to control for study variables while protecting safety and confidentiality that may have been compromised in the naturalistic setting.

**Data Analysis and Interpretation for Aim 1 and Aim 2**

Data analysis occurred after collection and cleaning of the data. PASW/SPSS Statistics 21 was the statistical analysis package used for data analysis. The investigator validated research assistant (RA) data abstraction, preparation, and cleaning.

**Preliminary data analysis.** Descriptive statistics were used to summarize (a) participant characteristics (e.g., age, gender, years of nursing experience, nursing education level, frequency of PCA use, PCA programming experience, work hours per week); and (b) measures of major study variables. Continuous variables were assessed for skew. Internal consistency for study measures were evaluated using Cronbach's alpha.

**Principal data analysis.** Study hypotheses were tested using the inferential statistical procedure, repeated measures-analysis of variance (RM-ANOVA). This study had one independent variable (IV) with three levels, three continuous dependent variables (EF1, CW, SF), and one nominal (A1) dependent variable. First, data was screened for missing values and outliers. Frequency distribution, histograms, and stem and leaf plots were examined for outliers. Preliminary dependent/paired t-tests were conducted to ensure repeated measure conditions are not significantly different. Second, to meet the assumptions for RM-ANOVA, continuous DV variables (EF1, CW, SF) were assessed for univariate normality and homogeneity of variance. Normality of each DV was confirmed. Assumptions of compound symmetry was confirmed with Pearson *r* for each DV variable set. Variation were equal across dependent variables to meet the assumption of compound symmetry. RM-ANOVA may inflate Type I error rate (Mertler & Vannatta, 2005). To reduce probability of Type 1 error, Bonferroni correction was performed to lessen the chance of Type 1 errors for dependent sample t-tests. Third, to compare the effects of the study independent variable (IV) to one nominal (A1) dependent variable McNemar test was

conducted. Should measures fail to meet assumptions' for ANOVA-RM, Friedman test will be used as the non-parametric alternative to RM-ANOVA. Post-hoc tests were used to identify significant differences between independent variables.

**Data Analysis and Interpretation for Aim 3**

To determine nurses' perceptions of the impact of interruption frequency on nurses' PCA interactions, qualitative content analysis was used for analysis, with phrases being the unit of analysis. This study used inductive content analysis (Elo & Kyngas, 2008). Data were coded as analysis began to build a model representing the data. Data that did not fit developing categories were coded to create new concepts. A data matrix was created to reanalyze sections of text as emerging results provided new insights.

Results will be reported by describing categories and sub-categories consistent with the data analysis method. Description of content using actual phrases will aid in the description or the study results. Inductive content analysis results may lead to modification of the proposed conceptual model for a study or to the development of a new model (Elo & Kyngas, 2008). Content analysis findings will be validated by a panel of nurses and will be used to explain finding from quantitative study.

**Trustworthiness and rigor.** The qualitative approach to this study applied multiple methods to improve trustworthiness, quality, and rigor (Elo & Kyngas, 2008; Lazar et al., 2010; Lincoln & Guba, 1985). The incorporation of qualitative analytic software into the study improves dependability. Reflexivity was maintained as the researcher balanced sensitivity with prior experience and bias during data collection, analysis, and interpretation for results. The researcher consulted with experts when appropriate to improve trustworthiness of data, such as

the review of content analysis results. Methods have been described and layered to improve trustworthiness, quality, and rigor throughout the study proactively highlight the strengths of the study and offset known limitations.

## Expected Findings/Interpretation of Results

Main effects will be reported with *F* statistics as significant or not significant.

- *Hypothesis #1:* Increased frequency of interruption will have a negative effect on nurses' performance efficiency (EF1-task completion time) and effectiveness (A1-accuracy). The results of the RM-ANOVA will assess within subject changes in efficiency (EF1). The McNemar test will be used to report the difference in effectiveness (A1) between conditions.
- *Hypothesis #2.* Increased interruption frequency will decrease nurses' subjective satisfaction and increase subjective workload. The results of the RM-ANOVA will assess within subject changes in satisfaction and subjective workload (NASA-TLX) between conditions.

## Summary

Chapter 3 described the methodological approach to the study. The population, setting, and sample were described. The measures and their reliability and validity were discussed. Data collection procedures were detailed and plans for data analysis were described. The following chapter will report the results of the study in Chapter 4.

## CHAPTER FOUR: RESULTS

The purpose of this study was twofold: (1) to quantify the impact of interruption frequency on registered nurses' performance, satisfaction, and subjective workload during PCA interaction and (2) to determine nurses' perceptions of the impact of interruption frequency. Study findings are presented in this chapter. First, this chapter presents the descriptive statistics. Next, the results of the preliminary data analysis are presented. Finally, the principle data analyses for each of the three research questions are described.

### Descriptive Statistics

A total of nine participants took part in the study. All participants were female. Each participant was exposed to all four levels of the independent variable, interruption frequency. Interruption frequency was categorized into four levels: condition A, condition B, condition C, and condition D. In condition A, participants completed tasks in an interruption free environment. Participants were exposed to two planted interruptions per 10-minute task scenario in condition B, and four planted interruptions per 10-minute task scenario in condition C. Finally, condition D was comprised of six planted interruptions per 10-minute task scenario.

The majority of participants did not require glasses or contacts (7, 78%), and none of the participants were colorblind. Slightly less than half of participants were white (4, 44%). The highest level of educational degree in nursing at the time of study was a bachelor's degree (5, 56%); four participants (44%) held an associate degree in nursing at the time of the study. Table 9 presents the frequencies and percentages for participant demographics.

Table 9. Frequencies and percentages for participant demographic information.

| | ***N*** | **%** |
|---|---|---|
| Gender | | |
| Female | 9 | 100 |
| Vision | | |
| I am farsighted, and I wear glasses or contact lenses. | 1 | 11 |
| I am nearsighted, and I wear glasses or contact lenses. | 1 | 11 |
| I do not wear glasses or contacts. | 7 | 78 |
| Colorblindness | | |
| I am not colorblind. | 9 | 100 |
| Ethnicity | | |
| Black or African American | 3 | 33 |
| Hispanic or Latino | 2 | 22 |
| White | 4 | 44 |
| Highest Level of Nursing Education | | |
| Associate Degree in Nursing | 4 | 44 |
| Baccalaureate Degree in Nursing | 5 | 56 |

Participants ranged in age from 27 to 46, with a mean age of 36 (*SD* = 6.76). Years of experience practicing as a registered nurse spanned 1 to 19 years for participants, with a mean of 6 years (*SD* = 5.46). On average participants were employed at their current hospital 6 years (*M* = 5.89, *SD* = 4.80); participants' years of experience on their current surgical unit ranged from 1 to 14 with an average length of 5 years (*M* = 5.22, *SD* = 4.35). Table 10 presents the means and standard deviations.

Table 10. Means and standard deviations for participant demographic information.

| | Minimum | Maximum | M | SD |
|---|---|---|---|---|
| Age | 27 | 46 | 36.22 | 6.76 |
| Indicate the number of years that you have been practicing as a registered nurse. | 1.0 | 19.0 | 5.94 | 5.46 |
| How many years you have worked at this hospital institution? | 1.0 | 14.0 | 5.89 | 4.80 |
| How many years have you worked on this medical-surgical unit? | 1.0 | 14.0 | 5.22 | 4.35 |

Participating nurses typically worked 36–40 hours per week (7, 78%) at hospitals with 300–399 beds (5, 56%). Two (22%) participants held specialty certifications from a professional organization; their certifications were Certified Lactation Consultant/Maternal Newborn Nursing (CLC/MNN) and Registered Nurse-Certified in Maternal Newborn Nursing (RN-MNN). The majority of participants worked in units that were a combination of medical and surgical patients (6, 67%). The type of specialty unit in which participants were employed varied within the study, with four (44%) participants working on units specifically for women such as mother-baby, OB/GYN, and post-partum caesarian section units. Table 11 presents frequencies and percentages for participants' professional experience information.

Table 11. Frequencies and percentages for participants' professional experience information.

| | *N* | % |
|---|---|---|
| Type of Unit | | |
| Combination of medical and surgical patients | 6 | 67 |
| Medical patients only | 2 | 22 |
| Surgical patients only | 1 | 11 |
| Specialty of Medical-Surgical Unit | | |
| Mother-Baby | 2 | 22 |
| Cardiac | 1 | 11 |
| Cardiac-Vascular-Neuro | 1 | 11 |
| Neurology | 1 | 11 |
| OB/GYN | 1 | 11 |
| Oncology | 1 | 11 |
| Post-Partum C-Section Unit | 1 | 11 |
| Telemetry | 1 | 11 |
| Certification in Specialty Area | | |
| No | 7 | 78 |
| Yes | 2 | 22 |
| Specialty Certification | | |
| None | 7 | 78 |
| CLC/MNN | 1 | 11 |
| RN-MNN | 1 | 11 |

All participants reported feeling very comfortable with using a computer, and the majority of respondents reported accessing the internet once or more per day (7, 78%). Six participants reported feeling very comfortable with the use of a PCA device; almost half reported using a PCA device a few times a month in their workplace (4, 44%). More than half of the participants (5, 56%) had no experience within their nursing practice with the pump employed in the study, the Baxter PCA II Pump; five (4, 56%) of the participants' hospital of employment had ANCC Magnet Status. Regarding the overall effect of interruption frequency, five (56%) participants reported low impact of interruptions on their performance. Five (56%) participants reported a low impact of interruption frequency on their satisfaction. Participants were evenly split between low, moderate, and high impact on the overall impact of interruption frequency on

their subjective workload (3, 33%). Table 12 presents frequencies and percentages for technology use, technology comfort, ANCC magnet status, and impact of interruption frequency.

Table 12. Frequencies and percentages for technology use, technology comfort, ANCC magnet status, and impact of interruption frequency.

| | *N* | % |
|---|---|---|
| Comfort with Computers | | |
| Very Comfortable | 9 | 100 |
| Internet Use | | |
| A Few Times in a Week | 2 | 22 |
| Once or More a Day | 7 | 78 |
| Comfort with PCA | | |
| Not Very Comfortable | 1 | 11 |
| Somewhat Comfortable | 2 | 22 |
| Very Comfortable | 6 | 67 |
| PCA Use | | |
| A Few Times a Month | 4 | 44 |
| A Few Times a Week | 3 | 33 |
| Once or More a Day | 2 | 22 |
| Baxter Pump Use | | |
| No | 5 | 56 |
| Yes | 4 | 44 |
| ANCC Magnet Status | | |
| No | 4 | 44 |
| Yes | 5 | 56 |
| Impact of Interruption Frequency on Performance | | |
| Low Impact | 5 | 56 |
| Moderate Impact | 4 | 44 |
| Impact of Interruption Frequency on Satisfaction | | |
| No Impact | 1 | 11 |
| Low Impact | 5 | 56 |
| Moderate Impact | 2 | 22 |
| High Impact | 1 | 11 |
| Impact of Interruption Frequency on Subjective Workload | | |
| Low Impact | 3 | 33 |
| Moderate Impact | 3 | 33 |
| High Impact | 3 | 33 |

**Descriptive Statistics for Main Measures**

For satisfaction scores, measured as frustration on the NASA-TLX, condition A ($M$ = 2.22, $SD$ = .972) had the lowest range, 1.0 to 4.0; condition C ($M$ = 6.44, $SD$ = 5.03) had the highest range, 2.0 to 15.0. Subjective workload scores (raw) for condition A ($M$ = 23.00, $SD$ = 10.87) and condition B ($M$ = 26.00, $SD$ = 11.14) ranged from 12.00 to 47.00; the range was highest for condition C ($M$ = 31.56, $SD$ = 22.31) at 12.00 to 78.00. For the NASA-TLX sub-scales of mental demand and temporal demand, there was a trend of increasing mean mental demand and increasing mean temporal demand as the amount of interruption increase from condition A to condition D. Total number of errors for participants ranged from 0.00 to 6.00 ($M$ = 1.56, $SD$ = 1.88). Time to complete tasks, recorded in seconds, ranged from 189.00 to 419.00 ($M$ = 292.11, $SD$ = 73.25) for condition B. The range for condition A was highest at 283.00 to 544.00 ($M$ = 385.67, $SD$ = 94.71). Table 13 presents the means and standard deviations for this data.

Table 13. Means and standard deviations for frustration scores, subjective workload scores (raw score with subscales), and efficiency (in seconds).

| | *Minimum* | *Maximum* | *M* | *SD* |
|---|---|---|---|---|
| Frustration A | 1.0 | 4.0 | 2.22 | .972 |
| Frustration B | 2.0 | 7.0 | 3.56 | 1.81 |
| Frustration C | 2.0 | 15.0 | 6.44 | 5.03 |
| Frustration D | 2.0 | 11.0 | 6.11 | 3.69 |
| Subjective Workload Raw Score A | 12 | 47 | 23.0 | 10.9 |
| Mental Demand A | 1.0 | 16.0 | 6.11 | 4.73 |
| Physical Demand A | 1.0 | 6.0 | 3.11 | 1.62 |
| Temporal Demand A | 2.0 | 7.0 | 3.67 | 1.73 |
| Performance A | 2.0 | 5.0 | 3.11 | 1.17 |
| Effort A | 1.0 | 11.0 | 4.78 | 3.35 |
| Frustration A | 1.0 | 4.0 | 2.22 | .972 |

| | *Minimum* | *Maximum* | *M* | *SD* |
|---|---|---|---|---|
| Subjective Workload Raw Score B | 12 | 47 | 26.0 | 11.1 |
| Mental Demand B | 2.0 | 11.0 | 6.33 | 3.35 |
| Physical Demand B | 2.0 | 8.0 | 3.89 | 2.42 |
| Temporal Demand B | 2.0 | 10.0 | 4.67 | 2.50 |
| Performance B | 1.0 | 8.0 | 3.22 | 2.11 |
| Effort B | 2.0 | 9.0 | 4.33 | 2.29 |
| Frustration B | 2.0 | 7.0 | 3.56 | 1.81 |
| Subjective Workload Raw Score C | 12 | 78 | 31.6 | 22.3 |
| Mental Demand C | 2.0 | 15.0 | 6.44 | 4.25 |
| Physical Demand C | 2.0 | 4.0 | 2.44 | .727 |
| Temporal Demand C | 2.0 | 15.0 | 5.33 | 4.85 |
| Performance C | 1.0 | 18.0 | 5.00 | 5.39 |
| Effort C | 2.0 | 12.0 | 5.89 | 4.51 |
| Frustration C | 2.0 | 15.0 | 6.44 | 5.03 |
| Subjective Workload Raw Score D | 13 | 68 | 36.7 | 20.3 |
| Mental Demand D | 2.0 | 16.0 | 7.67 | 5.05 |
| Physical Demand D | 2.0 | 11.0 | 5.11 | 4.01 |
| Temporal Demand D | 2.0 | 13.0 | 5.78 | 3.77 |
| Performance D | 2.0 | 9.0 | 4.78 | 2.59 |
| Effort D | 2.0 | 15.0 | 7.22 | 4.52 |
| Frustration D | 2.0 | 11.0 | 6.11 | 3.69 |
| Condition A Total Task time | 283 | 544 | 385.67 | 94.7 |
| Condition B Total Task Time | 189 | 419 | 292.11 | 73.2 |
| Condition C Total Task Time | 253 | 427 | 351.67 | 67.3 |
| Condition D Total Task Time | 291 | 507 | 396.44 | 74.7 |

**Preliminary Data Analysis**

Univariate data were screened for outliers using standardized values, or *z* scores. Values below -3.29 or above 3.29 were to be treated as outliers and removed from the dataset (Tabachnik & Fidell, 2012). No outliers were found in the univariate data, therefore no data were removed. Scores for frustration, subjective workload, and efficiency, measured as condition total task time, were tested for normality using boxplots. Boxplots for these three

variables revealed violations to the assumption of normality. Reliability testing was conducted on the six items that composed the subjective workload composite score for each condition to establish reliability. Reliability determines if the scores computed by the survey instrument are useful and significant; or in other words, reliable. The Cronbach's alpha test of reliability provides mean correlation between each pair of items and the number of items in a scale as alpha coefficients (Brace, Kemp, & Snelgar, 2006). According to the rules of thumbs (George & Mallery, 2010), alpha coefficients range from unacceptable to excellent where > .9 is Excellent, > .8 is Good, > .7 is Acceptable, > .6 is Questionable, > .5 is Poor, < .4 is Unacceptable. The subjective workload composite score with the highest alpha coefficient ($\alpha = .92$) were conditions C and D, indicating excellent reliability. The subjective workload score for condition A had the lowest alpha coefficient ($\alpha = .78$), indicating an acceptable reliability. Table 14 presents the alpha coefficients.

Table 14. Cronbach alpha reliability for subjective workload composite scores.

| Score | Items | Cronbach α |
|---|---|---|
| Raw Score A | 6 | .78 |
| Raw Score B | 6 | .85 |
| Raw Score C | 6 | .92 |
| Raw Score D | 6 | .92 |

## Data Analysis

### RQ1: What is the Effect of Interruption Frequency on Performance Efficiency and Effectiveness of Medical-Surgical Nurses' PCA Use?

*$H_0$1: Increased frequency of interruption will have no effect on nurses' performance efficiency and effectiveness.*

To assess Research Question One, the researcher conducted a Friedman test to investigate the effect of interruption frequency on efficiency, measured as participants' total task time measured in seconds for each condition type, and a McNemar's Test to test the effect of interruption frequency on effectiveness. Effectiveness was measured as accuracy within the study; participants who committed no errors were accurate, or effective, and those who committed errors were inaccurate. Because nominal data was gathered for this variable, the researcher employed McNemar testing. Wilcoxon testing was conducted for pairwise comparisons of efficiency by condition.

In preliminary analysis, the assumption of normality was assessed for efficiency, measured as total task time per condition. The assumption of normality was violated; therefore, the researcher conducted the nonparametric alternative to the repeated measures ANOVA, the Friedman test. Results of the Friedman test were significant for the main effect of interruption frequency on efficiency $\chi^2(3) = 14.60$, $p = .002$, suggesting that interruption frequency affected condition total task time for participants (see Table 15).

Table 15. Results of the Friedman test for main effect impact on efficiency by condition type.

| ***N*** | ***X*²** | ***df*** | ***p*** |
|---|---|---|---|
| 9 | 14.6 | 3 | .002 |

Pairwise comparisons were conducted using the Wilcoxon Signed Rank test to determine significant differences in the mean of total task time by condition type. There was a significant difference in mean time between conditions A and B, and B and D ($p < .05$). The mean time for condition A ($M = 385.67$, $SD = 94.71$) and mean time for condition D ($M = 396.44$, $SD = 74.69$)

were both higher than the mean time for condition B ($M$ = 292.11, $SD$ = 73.25). Table 8 presents the results of the pairwise comparisons.

Table 16. Results for the Wilcoxon signed rank test pairwise comparisons for efficiency by condition type.

| | ***N*** | ***Z*** | ***p*** |
|---|---|---|---|
| Condition A-B | 9 | -2.67 | .008 |
| Condition A-C | 9 | -1.48 | .139 |
| Condition A-D | 9 | -.533 | .594 |
| Condition B-C | 9 | -1.72 | .086 |
| Condition B-D | 9 | -2.67 | .008 |
| Condition C-D | 9 | -1.72 | .086 |

To assess the effect of interruption frequency on effectiveness, a McNemar test was conducted. The nominal variable in this analysis was effectiveness. Participants were considered effective if they were able to assist with PCA use without errors. Results of the six comparisons included in the McNemar test did not show significance, suggesting there was not an impact of interruption frequency on effectiveness (see Table 17).

Table 17. Results of the McNemar test for effectiveness by condition type.

| | **Cond. A-B** | **Cond. A-C** | **Cond. A-D** | **Cond. B-C** | **Cond. B-D** | **Cond. C-D** |
|---|---|---|---|---|---|---|
| *N* | 9 | 9 | 9 | 9 | 9 | 9 |
| *p* | 1.00 | .125 | 1.00 | .125 | 1.00 | .063 |

**RQ2: What is the Impact of Interruption Frequency during PCA Interactions on Medical-Surgical Nurses' Perceptions of Satisfaction and Subjective Workload?**

*$H_0$1: Increased frequency of interruption will have no effect on nurses' perceptions of satisfaction and subjective workload.*

To assess Research Question Two, the researcher conducted Friedman testing to test the impact of interruption frequency on satisfaction, as measured by participants' frustration scores per condition type on the NASA-TLX, and to test the impact of interruption frequency on subjective workload, as measured by participants' raw score per condition type on the NASA-TLX. Friedman tests were conducted for pairwise comparisons.

In the preliminary analysis for the repeated measures ANOVA for the impact of interruption frequency on participants' satisfaction scores, the results of the Mauchly's test for Sphericity showed that the assumption of sphericity was violated ($p < .05$). The researcher conducted Friedman testing to investigate the impact of interruption frequency on satisfaction scores since this assumption was violated. Significance was found for the main effect of impact of interruption frequency to satisfaction scores (see Table 18).

Table 18. Results of the Friedman test for main effect impact of satisfaction score by condition type.

| ***N*** | ***$X^2$*** | ***df*** | ***p*** |
|---|---|---|---|
| 9 | 9.47 | 3 | .024 |

Wilcoxon signed rank tests were conducted for pairwise comparisons between satisfaction scores by condition type. Significance was found for frustration score comparisons between A–D, B–C, and B–D ($p < .05$). Mean frustration score for condition D ($M = 6.11$, $SD$ =

3.69) was significantly higher than both mean frustration scores for condition A ($M$ = 2.22, $SD$ = .972) and condition B ($M$ = 3.56, $SD$ = 1.81). The mean of participants' frustration score was higher for condition C ($M$ = 6.44, $SD$ = 5.03) than for condition B ($M$ = 3.56, $SD$ = 1.81). Participants in condition D reported higher frustration scores than participants in conditions A and B. This result reflects a trend of increasing frustration as the amount of interruptions increased (Figure 8). Table 19 presents results for the pairwise comparisons.

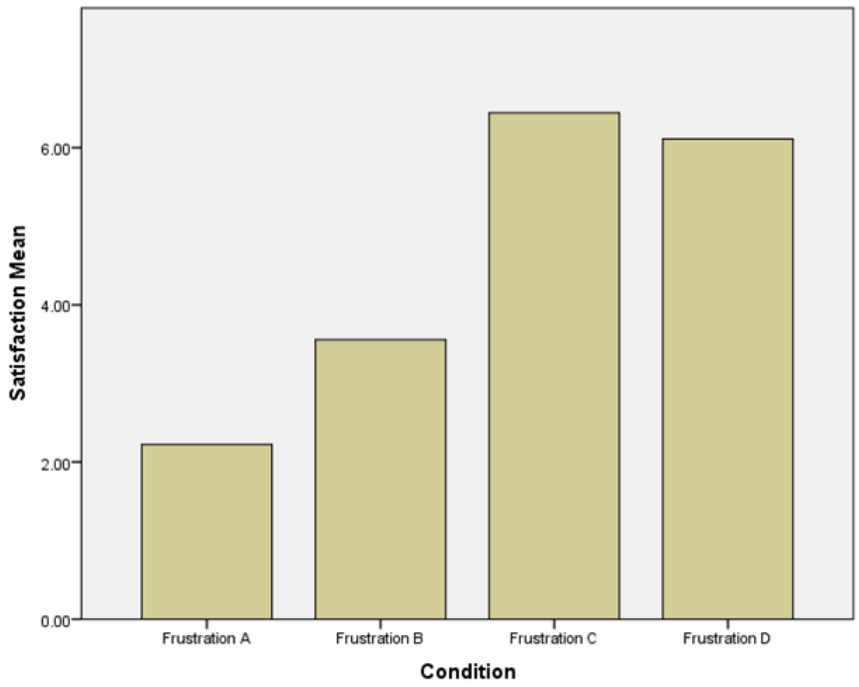



Figure 8. Trend of frustration scores by condition.

Table 19. Results for the Wilcoxon signed rank test pairwise comparisons for satisfaction score by condition type.

| | *N* | *Z* | *p* |
|---|---|---|---|
| Frustration A-B | 9 | -1.70 | .088 |
| Frustration A-C | 9 | -1.96 | .050 |
| Frustration A-D | 9 | -2.20 | .028 |
| Frustration B-C | 9 | -2.00 | .045 |
| Frustration B-D | 9 | -2.21 | .027 |
| Frustration C-D | 9 | -.281 | .779 |

To assess the impact of interruption frequency on participants' subjective workload scores, a Friedman test was conducted. In preliminary analysis, it was determined that the subjective workload scores violated the assumption of normality; therefore, a nonparametric test was conducted to compare subjective workload scores. Friedman testing showed no significance for the main effect impact of interruption frequency on subjective workload score by condition type (see Table 20). These results suggest that participants have no differences in perceptions of subjective workload by condition type, as measured with the NASA-TLX. However, the data demonstrate a trend of increasing subjective workload (raw) as the amount of interruptions increased (Figure 9).

Table 20. Results of the Friedman test for main effect impact of subjective workload score by condition type.

| *N* | $X^2$ | *df* | *p* |
|---|---|---|---|
| 9 | 1.88 | 3 | .599 |

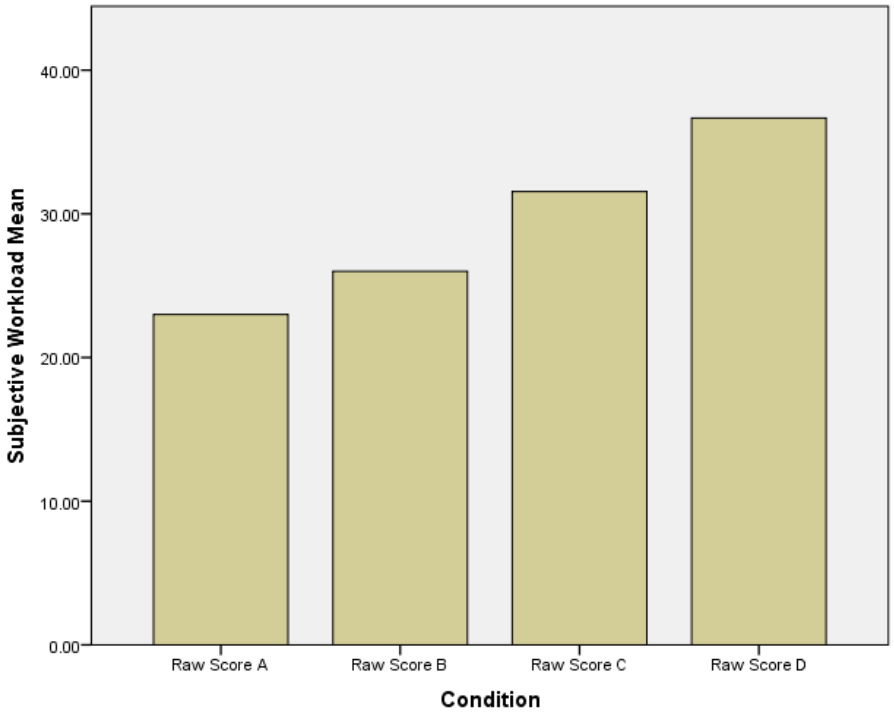



Figure 9. Trend of increasing subjective workload (mean) by condition.

**RQ3: What are Medical-Surgical Nurses' Perceptions of the Impact of Interruption Frequency during PCA Interactions?**

The assess Research Question Three, the transcribed contents of eight semi-structured interviews were analyzed within Dedoose® using inductive qualitative content analysis (Elo & Kyngas, 2008). During the initial stage of analysis, the transcripts were read numerous times to get an overall impression of each participant's perceptions and responses. Next, coding was performed on each interview, whereby similar content was grouped together into categories. The unit of analysis was phrases. Categories were derived inductively during the process of analysis. After each interview was coded, all transcript content and categories were reviewed again for accuracy of coding and completeness of categories. Coding of the manifest content resulted in 21 sub-categories that described nurses' perceptions of the impact of interruption frequency during

PCA interactions (see Table 21). Next, new categories were developed and old categories were revised then grouped into higher order headings (Elo & Kyngas, 2008). Higher order groupings reduced the number of categories allowing for abstraction of the data into the generic descriptive categories. Analysis of the semi-structured interviews resulted in the identification of two main categories describing medical-surgical nurses' perceptions of the impact of interruption frequency during PCA interactions: the *nature of interruptions* and *nurses' reaction to the interrupted work environment*. The following section provides is a description of the main category, generic, and sub-category abstraction results (see Figure 10) using participants' statements to illustrate each category.

Table 21. Initial inductively derived sub-categories.

| | |
|---|---|
| Auditory interruptions | Feeling the need to stop and start |
| Background noises | Feeling worried |
| Becoming accustomed to the environment | Frequency of interruptions |
| Checking for mistakes | Maintaining focus |
| Double-checking work | Maintaining patient safety |
| Feeling annoyed | Multi-tasking |
| Feeling anxious | Physical symptoms of stress |
| Feeling frustration | Slowing down |
| Feeling hurried | Trying not to forget |
| Feeling timed | Visual interruptions |
| Feeling the need to slow down | |

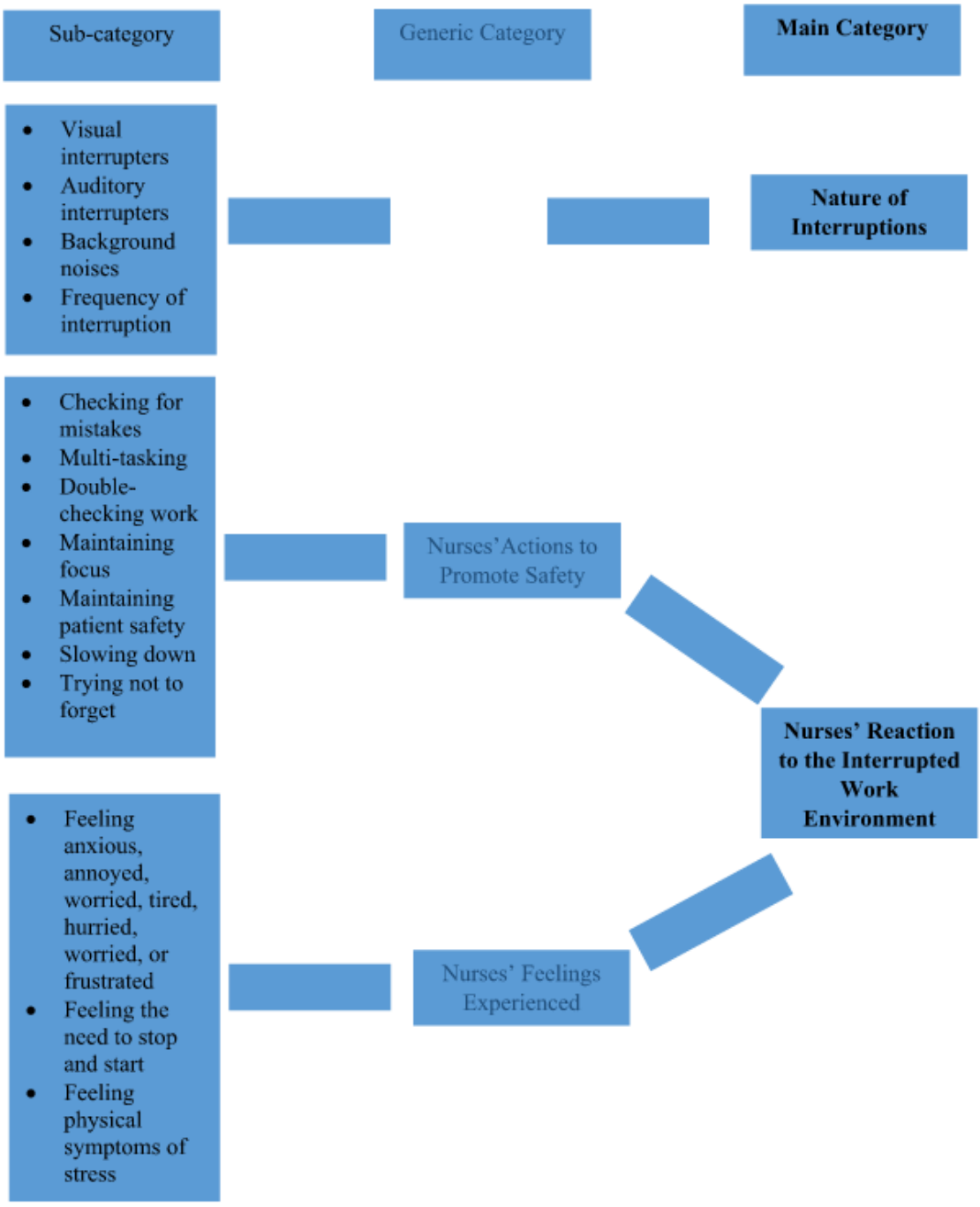



Figure 10. Category abstraction results.

**The Nature of Interruptions**

Interruption frequency impacts the overall state of the medical-surgical nurses' work environment. Nurses described the overall interruption-laden conditions of their work environment through their descriptions of the interruptions they experience each day. This main category was described by one main category: *the nature of interruptions*.

The inherent features of interruption experienced by the medical-surgical nurses were described as visual interruptions, auditory interruptions, background noise, and frequency of interruptions. *Visual interruptions* were described by the nurses as their awareness that the appearance or actions of a patient or setting drew the nurses' attention away from PCA related tasks.

> "Obviously if the patient doesn't look good, if they're pale or don't look right, that will make me stop."
>
> "I know this sounds crazy but a lot of clutter in the room will make me stop what I am doing."
>
> "Visual, still would go back to the television…your attention might turn to that."
>
> "Lot of times I have patients that are showing me things … as I'm doing a task."

*Auditory interruptions* were described by nurses hearings a patient-related or setting-related sound that drew the nurses' attention away from PCA or related tasks.

> "Yeah the phone interrupts me a lot. Um, or I guess IV pumps, bed alarms, those are things I would be hearing."
>
> "The auditory is definitely almost constant … phone ringing, and bed monitors."
>
> "Um, again with patients interrupting with conversation."

Nurses described *background noise* as the sounds from sources such as technology or other persons that are constant at times and that need to be ignored or tuned out. Background noise does not interrupt nurses' work.

> "I try to do the same thing, I try to focus on what I'm doing so if the TV's blasting in the patient's room, if it's really blasting and it's annoying me I might turn it down, but for the most part I just try to tune it out."
>
> "A lot of background noise. TVs, babies crying."
>
> "The hospital has a lot of older people so ... they're always super loud so I tend to block them out."

Nurses used a broad range of terms to describe the *frequency of interruptions* experienced in their medical-surgical work environment.

> "Constantly."
>
> "It's sporadic."
>
> "You never know when you're going to be interrupted."
>
> "So I'm constantly interrupted with what I'm doing."
>
> "Very frequent."
>
> "Okay, so we are constantly interrupted."

**Nurses' Reaction to the Interrupted Work Environment**

This main category was described by two generic categories. A *reaction* is an action performed or a feeling experienced in response to a situation or event (Oxford Dictionaries, 2014). Nurses described actions and feelings in response to the frequently interrupted work environment.

**Nurses' actions to promote safety**. In response the interruption-laden work environment, nurses react by taking specific actions that promote patient safety. Nurses described *maintaining focus*.

"To get to the other tasks but still trying to focus on what you're doing in the present moment."

"Concentrate on what you're supposed to do right now."

"I try to focus on what I'm doing."

"That's a constant, constant thing. Um, because it's harder to like concentrate I mean sometimes you have to add the prescription to things, you've got to be right on point. So it's harder to like focus.

*Multi-tasking* was described by nurses as an action required as a results of being frequently interrupted in the work environment.

"Sometimes you have to talk to them. That's another interrupter that you have to actually respond to them."

"Um, it's not an issue because if I'm interrupted I will repeat the process and re-verify it. It may take me longer to re-program the PCA, but you have to focus on what you're doing so if I'm interrupted I'll start over."

"Lots of questions from patients while I'm doing a task."

Nurses described *double checking their work* in the interruption-laden work environment.

"Double, triple checking my work because I'm worried I'm going to make a mistake."

"Um, it could be frustrating at times I have to remind myself to take my time and also make sure to get it double verified behind me just to make sure there's no mistakes."

"It's independently double verified and we double verify the medication."

Other actions that nurses described to promote patient safety included *becoming accustomed to the environment*, *trying not to forget*, *checking for mistakes*, and *slowing down*.

**Nurses' feelings experienced**. In response the interruption-laden environment, nurses experience feelings that impacted their state of mind while performing PCA and related tasks. Most commonly, nurses feel anxious, frustrated, and hurried.

> "It makes me anxious. Very nervous. I tend to double check myself more when I'm interrupted, because I'm worried I'm going to put in the wrong settings."
>
> "It makes me anxious, nervous, that's why I always like to ask someone to double check and that's why we have independent verification at work, because of that.
>
> "Um, it could be frustrating at times I have to remind myself to take my time and also make sure to get it double verified behind me just to make sure there's no mistakes."
>
> "It can cause frustration which makes it harder to concentrate on what you're doing. Or you've got other things on your mind. Other things you're thinking or doing."
>
> "I feel like I have to be a little bit more hurried."

Nurses also described how the feel symptoms of stress including, "chest pressure and pain" and "increase in heart rate, maybe a little sweaty."

## Summary

Chapter Four presented the results of the preliminary and principle data analyses for research questions one, two, and three. Chapter Five will present a discussion of the results as well as the study limitations and implications for future research.

# CHAPTER FIVE: DISCUSSION

The aims of this study were (1) to determine the impact of interruptions frequency on nurses' PCA performance; (2) to determine the impact of interruptions frequency after PCA interactions on nurses' subjective satisfaction and subjective workload; and, (3) to determine nurses' perceptions of the impact of interruptions frequency on nurses' PCA interactions. The central hypothesis of this study was that interruption frequency during nurses' patient-controlled analgesia (PCA) device interaction will affect nurses' performance efficiency and effectiveness, subjective satisfaction, and perceived subjective workload. A mixed-method approach was used. First, an experimental repeated-measures design was used to quantify the impact of interruption frequency for aims one and two. After each experiment, semi-structured interviews were used to collect data that were analyzed to determine nurses' perceptions of the impact of interruption frequency for aim three.

The following chapter discusses the results of the study. This chapter first provides a summary and interpretation of the results. The context of the results are then discussed in comparison the current literature. The theoretical, methodological, and nursing implications as well as limitations of the study are addressed. This chapter concludes with suggestions for future research.

## Summary and Interpretation of Study Results

### Research Question One

Research question one asked, what is the effect of interruption frequency on the efficiency and effectiveness of medical-surgical nurses' PCA use? Hypothesis one stated that increased frequency of interruption will have a negative effect on nurses' performance efficiency

(EF1-task completion time) and effectiveness (A1-accuracy). The findings for hypothesis one were partially supported. First, this study found that increased frequency of interruption had a significant negative effect on efficiency as measured in total task time (seconds), as hypothesized. Similar results have previously been reported. Using a direct observation method in the naturalistic setting, Trbovich, Prakash, Stewart, Trip, and Savage (2010) reported that mean total task times for registered nurses' general medication administration interrupted tasks were significantly longer ($F1,11 = 101$, $p < .001$) than non-interrupted tasks. The study by Trbovich, et al., quantified total task time for interruptions during general medication administration while the current dissertation study honed in on the high-risk PCA programming tasks. Knowledge regarding nurses' performance efficiency in both situations is important because interrupted work negatively impacts the nurses' ability to adequately meet patients' care needs (Trbovich, et al.). Thus, the current study supports and extends what is known about nurses' performance efficiency as measured in total task time.

Second, this study found that increased frequency of interruption did not have a significant effect on effectiveness, measured as accuracy. Although the finding was not statistically significant, this finding is clinically significant for two reasons. First, only two of nine nurses committed no errors in this study; seven nurses committed a total of 15 errors while programming the PCA. Of the 15 errors committed, most ($n$=10) occurred immediately after being interrupted during condition C in which nurses experienced four interruptions and during condition D in which nurses experienced six interruptions, suggesting that interruption does impact effectiveness. Using a direct observation method, Westbook et al., (2010) reported that interruption frequency was significantly associated with medication administration errors; the more interruptions nurses experienced, the greater the number of medication errors. Specifically,

the risk of an error doubled in the presence of four or more interruptions in this study. Other research suggested that interruptions of 2.8 seconds could double the error rate, and interruptions of 4.4 seconds could triple the errors rate (Altman, Trafton, and Hambrick, 2013). The current study contributes a new perspective on nurses' PCA programming accuracy that should be further studied.

The second reason for clinical significance is that five errors that occurred were in no immediate relation to the planted interruptions within this study; three of these five errors occurred in condition A in which there were no planted interruptions. Based on this study, is not possible to assign a specific reason for each of these five errors that occurred; however, the errors may be explained based on what is known about effectiveness (accuracy) from relevant current literature which reports that errors most frequently occur as a result of human factors, limited inexperience, or device issues (Hicks, et al, 2008) or potentially nervousness, especially when the errors occurred in condition A. Whether instigated as a result of an interruption or not, accurate completion of programming tasks is critical to safe administration of PCA and achieving patient safety. This study affirms our limited understanding of the nurse-device interaction and supports concerns regarding the impact of interruptions on nurses' PCA programming effectiveness (accuracy).

**Research Question Two**

Research question two asked, what is the effect of interruption frequency on medical-surgical nurses' subjective satisfaction and subjective workload with PCA use? Hypothesis two stated that increased interruption frequency will decrease nurses' subjective satisfaction and increase subjective workload. Hypothesis two was partially supported. First, satisfaction was

defined as attitudes toward the use of a system, with an attitude being a settled way of thinking or believing about someone or something (Oxford Dictionaries, 2015). This study found that increased interruption frequency significantly decreased nurses' subjective satisfaction, as hypothesized. Known current literature has not reported or quantified the effects of frequency of interruption on medical-surgical registered nurses' subjective satisfaction and thus, this finding adds new knowledge to current literature. This study demonstrated a trend of increasing frustration as the amount of interruptions increased. Nurses' satisfaction with medical devices describes positive and negative impact on nursing practice, such as increased risk of errors and overly complex programming tasks (Marini, Hasman, Huijer, & Dimassi, 2010). Knowledge regarding nurses' subjective satisfaction is important because it describes how users feel about a system or device, which links aspects of the human-device interaction, such as interface design, tasks, or environment that did not meet users' needs limitations, or expectation (Fairbanks, Caplan, Bishop, Marks, & Shah, 2007; Garmer, Liljegren, Osvalder, & Dahlman, 2002). User satisfaction provides valuable insight into the quality of an interaction and how users were impacted by device interactions.

Second, this study found that increased interruption frequency did not significantly increase nurses' subjective workload. Known current literature has not reported or quantified the effects of interruption frequency on medical-surgical registered nurses' subjective workload and thus, this finding adds new knowledge to current literature. Although not statistically significant, the subjective workload scores did increase incrementally from condition A to B, B to C, and then C to D. In related literature, only one study has addressed nurses' PCA use and subjective workload. Kataoka, Sasaki, and Kanda (2011) addressed the nurses' subjective workload during infusion pump use. They found that nurses experienced an increase in subjective workload due to

time pressure during shortened infusion pump operations. This finding from the current study indicates that additional study is needed regarding nurses' subjective workload and frequent interruption.

**Research Question Three**

Research question three asked, what are medical-surgical nurses' perceptions of the impact of interruption frequency during PCA interactions? This study found that two main categories (Figure 10) that described medical-surgical nurses' perceptions of the impact of interruption frequency during PCA interactions: *the nature of interruptions* and *nurses' reaction to the interrupted work environment*. The findings described the negative effect of frequent interruption on the work environment whereby nurses subsequently implement compensating strategies to counterbalance the impact of interruption in the workplace.

In the current study, nurses described the *nature of frequent interruption* as visual and auditory interrupters, background noises, and frequency of interruption. The nature of interruptions describes the work environment created in which nurses must complete safe patient care while enduring frequent interruption. Using direct observation, Biron (2009) described sources of interruption such as individual (e.g., healthcare professional, patients, families) and technical (e.g., equipment, alarms). Biron's descriptions are comparable with nurses' descriptions from the current study.

This study is the first to describe the frequency through which nurses perceive interruption, using nurses' own words, such as "constantly" and "very frequently." Nurses' own words demonstrate how they are acutely aware of the interruption environment, as well as the actual and potential impact of interruptions. Interruption rates have been quantified during

nurses' general work and reported in several studies. Biron (2009) reported that interruptions ranged from 0.8 to 41.8 events per hour. Westbook, et al., (2010) reported that the effects of interruption doubled when comparing zero interruptions to four interruptions for a single medication administration task. This study is also the first to employ semi-structured interviews to collect data regarding nurses' perceptions immediately after nurses experience frequent interruption. This unique perspective adds new knowledge to what is known about interruptions from the nurses' perspective, improving our understanding of the work environment.

The main category in this study, *nurses' reaction to the interrupted work environment* is described as nurses' actions to promote safety and as nurses' feelings experienced. Nurses perceive that patient safety is negatively impacted by frequent interruption and nurses experience negative intrapersonal consequences as a results of frequent interruption, that have the potential to negatively impact performance, satisfaction, and subjective workload. Evidence that nurses react to the interrupted workplace and take steps to promote safety is a new finding that has not yet been reported in the literature. This is the first study to report nurses' responses to frequent interruption in relation to their efforts to counter-balance the effects of interruption. This supports descriptions of nurses' cognitive processes. Nurses reported that as a result of frequent interruption, they often check for mistakes, double check their work, try to maintain focus and patient safety, and attempt to slow down their work as well as try not to forget their primary task focus.

A finding unique to this study is the description of nurses' feelings experienced as a result of frequent interruption; no known study reports similar findings. Nurses' perceived a variety of interpersonal consequences such as feeling annoyed, worried or frustrated, feeling the need to stop and start their work, and feeling physical symptoms of stress such as chest pain,

increased heart rate, and sweatiness. The potential exists the frequent interruption creates time pressure that negatively impact patient safety and the nurses' reaction to the interrupted work environment. These interpersonal and physical symptoms have the potential to impact personal performance, satisfaction, and subjective workload during the nurse-PCA interaction. The findings from research question three of this study are the first to describe nurses' perceptions of interruption frequency in relation to PCA interactions and thus contributes new knowledge to what is currently know about the nurse-PCA interaction.

**Theoretical, Methodological, and Practical Implications**

This study provided important and meaningful findings to advance our understanding of interruptions and provide the groundwork for future study. There are theoretical, methodologic, and practical implications of this study. As the basis for this study, a conceptual model was synthesized from existing conceptual and theoretical models. The systems model of clinician interaction with medical devices (SMCIMD) was developed for understanding the effects of interruptions frequency during nurses' PCA use and served as a framework for describing nurses' perceptions of interruption frequency during PCA interactions. The results of research questions one, two, and three, provide data for as feedback during system re-design.

This mixed-method approach provided a holistic understanding of the nurse-PCA interaction than in comparison to using one approach, either the quantitative or qualitative. Human factors analysis encourages multiple, complimentary approaches rather than a single approach during evaluation (FDA, 2011). The current study supports continued use of multiple, complimentary methods in the study on interruption frequency and the nurse-PCA interaction.

There are several implications for nursing practice. There is a dearth of research describing the work of medical-surgical registered nurses (RN) during device-interactions; this is the first study to focus attention to the high-risk, high consequence PCA-interaction. Knowledge that medical-surgical RNs performance efficiency and effectiveness, cognition, and satisfaction are negatively impacted is important to RNs, nurse leadership, and health care organizations who desire to improve patient safety during the nurse-PCA interaction. Medical-surgical RNs, nursing leadership, and healthcare organization need education regarding the impact of interruption on nurses' performance, cognition, and satisfaction. In the future, evidence–based practices should be developed, tested, and applied to mitigate the impact of interruption during PCA use on nursing practice and overall patient safety.

To improve medical device safety, it is critical to understand how a medical device will be used, including the nature of its users and the environment (FDA, 2011). This study suggests that the design of this specific PCA did not support the needs and limitations of medical-surgical RNs during frequently interrupted environment. Medical device manufactures and FDA regulators need awareness of the findings to improving medical device design, use, and risk assessment throughout the device life-cycle.

## Limitations of the Study

Several limitations exist in relation to this study. First, the within-groups design was limited because it can be difficult for the researcher to control for learning effects. Training was planned to be sufficient to allow time for all participants to become familiar with the device and tasks and measurement tool, specifically the NASA-TLX. However, four of nine participants had previous experience with the specific Baxter II PCA used in this study. It is unclear if previous

experience with the Baxter II PCA study device may have affected study outcomes or if training on the NASA-TLX was clear and sufficient. Next, the sample size ($n$=9) was small but adequate for the design and conducted at one location. Participants were recruited from a specific region of west central Florida, convenient to the study site. Further, all participants ($n$=9) were female. Participants from this study may not be representative of other medical-surgical registered nurse populations. These factors limit generalizability of findings.

During the nurses' task performance in the high-fidelity simulation laboratory, the principal investigator was directly present and support personnel were present via a one-way glass window, observing and recording nurses' task performance with audio-visual equipment. There is a risk that the presence of observers and audio-video technology may have influenced nurses' performance during observation. The potential exists that nurses may have modified their behavior from the naturalistic setting. Further, the PCA programming task and mental workload conditions in this were developed to simulated the naturalistic setting but study were artificial. The nature and frequency of planted interruptions and PCA programming tasks were potentially not representative of the naturalistic setting. Although these issues may have been reduced by the use of four testing conditions, these environmental factors may limit generalizability.

With regards to the qualitative aspects of this study, multiple methods were applied to improve trustworthiness while offsetting known limitations. Dependability was supported with the incorporation of qualitative analytic software. There is the potential that bias may have been introduced during data collection while conducting the semi-structured interviews by means of the researcher's body language, tone, or follow-up questions. Bias may also have been introduced during data analysis. Therefore, reflexivity was maintained as the researcher balanced personal sensitivity with prior experience and bias during data collection, analysis, and

interpretation. Data collection and analysis was conducted by one researcher. This may have improved consistency during data collection using the semi-structured interview but limited validity in the coding and analysis. A member check ($n$=2) of the results was completed with participants.

## Future Directions for Research

Future research should continue to consider medical-surgical nurses as a population of interest as there is little research describing medical-surgical nurse' practice and the nurse-PCA device interaction. Future studies should be expanded to represent medical-surgical nurses with a broader range of expertise and variability in individual characteristics.

The focus of this study was on the impact of frequency of interruption during medical-surgical nurses' interaction with PCA. However, during the completion of this study, it became clear that the timing of the interruption may have an impact on nurse-PCA interaction. Future studies should consider the possibility that the timing of interruptions may play a role in the nurse-PCA interaction as well as patient safety outcomes.

Only one model of PCA device was used in this study. Future study should be expanded to include different PCA models and potentially to compare the nurse-PCA interaction across various PCA models. An interesting perspective could be to establish PCA device compliance with established human factors user-centered design principles and then compare the nurse-PCA interaction based to user-centered design compliance.

This study was conducted in the simulation laboratory. However, additional study is needed to improve our understanding of nurses' perceptions of the impact of interruption. Qualitative methods may be useful to explore this area in the naturalistic setting.

Nurses in this study experienced a range of interpersonal and physical responses to frequent interruption. Future studies should explore nurses' interpersonal and physical response to frequent interruption. The potential exists the frequent interruption creates time pressure that negatively impact patient safety and the nurses' reaction to the interrupted work environment. These interpersonal and physical symptoms have the potential to impact personal performance, satisfaction, and subjective workload during the nurse-PCA interaction.

Patient safety continues to be a prime area of relevant research. The concentration on nurse interaction with medical devices and the evaluation of the user-device interaction are a fertile ground for a fundable, program of patient safety research. Nurses can be instrumental in developing, testing, and reporting for methodological improvements.

## Conclusion

The overarching purpose of this study was to improve our understanding of interruption frequency during medical-surgical nurses' PCA use. Prior to this study, little was known about the impact of interruption frequency on medical-surgical registered nurses' PCA interactions in relation to performance efficiency (task completion time) and effectiveness (accuracy), cognition, and subjective satisfaction. This is the first known empirical study both to quantify the effects of interruptions frequency on nurses' PCA interactions and to determine nurses' perceptions of the impact of interruption frequency. This study provided evidence that interruptions frequency negatively affects performance efficiency (task completion time) and subjective satisfaction (frustration). While not statistically significant, interruption frequency negatively impacted performance effectiveness (accuracy) and subjective workload as hypothesized. Nurses described the impact of interruption frequency in terms of negative impact

on the work environment and in terms of personal negative impact. The findings from this study improve our understanding of interruption as well as the nurse-PCA interaction and may subsequently be used to reduce PCA-related errors and improve patient safety.

**APPENDIX A**
**MEDICAL DEVICE USABILITY STUDIES**

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| Brixey, Zhang, Johnson & Turley (2009) | Descriptive, non-experimental [Level 3A] | -Registered nurses and physicians (*n*=19) -Infusion pump (*n*=1) dual channel volumetric -Infusion pump programming data monitoring, programming, -Naturalistic, intentive care unit | -Human factors contextul evaluation: 2 dual experts observed users interaction using the infusion pumps and measured ambiant light during a four-hour period. | -Device characteristics : screen measurement and font size, overall screen legibility -Environment: Ambiant light readings at various distances and locations with lights on and off | -Multiple factors identified including small font size, faint lighting, reduced screen contrast, and reduced legibility on infusion pump screen. -User work-around strategies: nurse workarounds included use of handmade tape labels attached to pump screen to improve clarity and enhanced legibility. -Environemnt: Pump position and re-positioning to view the screen contributed to interruptions in work flow leading to potential safety hazards. -Recommendations: Manufacturers should adherance to FDA recommendations to ensure legibility. An environmental approach is | -Reliability: interrater agreement for light measurement not reported. -Data collection during a single, 4-hour timeframe may not be representative of all lighting measures over a 24-hour day. |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| | | | | | recommended to determine microdisplay and small-screen devices in health care are legible and useful. | |
| Carayon et al., (2007) | Descriptive, non-experimental [Level 3A] | -Nurses -critical care and medical/surgical<br>-Bar code medication administration (BCMA) system<br>-Medication administration tasks<br>-Naturalistic setting: 472 bed academic hospital/acute care, | -Direct observation (*n*=62) medication administration in natural setting were conducted by a team of 2 (1-HFE expert and 1-pharmacist)<br>-Data collection on work-system model tool: task, BCMA system, organizational factors (interruptions), physical environment, | -Task characteristics -number of tasks, task sequence<br>-Device characteristics -Audible alarms, automation surprises<br>-Environment characteristics -Interruptions; patient room/isolation use; workflow. | -Tasks: 18 different task sequences identified, with broad variability steps sequence for medication administration processes; some sequences (*n*=10, 10%) included potentially unsafe acts.<br>-Device: Automation surprises (*n*=10, 10%) and audible alarms (*n*=26, 42%)<br>-Environment: Interruptions (*n*=20) observed; working conditions can hinder the medication administration process.<br>-Patient factors (e.g., isolation patients) made the BCMA-based tasks | -Limitations of structure observation relating to observer training, presence during data collection/task performance,<br>-Inter-rater reliability could not be reported<br>-Patient rooms under isolation precautions were not observed<br>-Timing of observations may have influenced type and number of interruptions<br>-Only 31% of nurses agreed to post observation interview |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| | | | individual nurses and patients<br>-Post observation interview | | during medication administration tasks difficult for nurses.<br>-Workflow and tasks changes are a result after introduction of new technology.<br>-Direct observation may be useful for identifying the work system factors that facilitate or hinder the tasks during medication administration leading to redesign to improve user efficiency, interaction with the technology, and patient safety.<br>-Conceptual framework was effective for system description. Work system model of patient safety (Carayon, et al, 2005). | |
| Carayon, Hundt, & Wetterneck | Descriptive [Level 3A] | -Nurses (*n*=600)<br>-Smart ® general | Longitudinal surveys of nurses' | -Questionaire for User Interface | -Participant description and response rates reported. | -Single source of data limit generalizability.<br>-QUIS was modified |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| (2010) | | infusion pump -Tasks-NA -Environment-academic hospital | expereince with impliemntation process and use of infusion pump: -pre-implementation survey -6-week-post-implementation survey -1-year-post-implementation survey | Satisfaction (QUIS®) User perceptions of: -Device implimentation process -Device performace -Device usability -User accpetance | -User perception of implimentation process: Nurses did not consistently improve from the pre- to post- survey; input into decision-making on pump implimnetation process did not consistently improve from pre- to the post-implementation survey; nurses perceived that they received more information before than after implimentation. Training at 6-week and 1 year post implimentation were more confusing. -Learnability: Nurses found pump somewhat positive and leanrin to use the pump became easier over one year. -Efficiency: Nurses perceived that the pump improved safety but | for this study to fit context of infusion pumps. Relaiblity of QUIS not established after questions were modified. -QUIS administere in paper and electronic form and over time. Effects of time and learning may have impactedmeasures. |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| | | | | | responses for ease of use durring an emergency were lower after one year. Signifgant findings after one year for nurses perceptions of efficiency to improve quality of are, accomplish tasks more efficiently, enhance effectiveness of my job, and increase safety of patients. -Satisfaction: Nurses preorted easier interaction s with the pump after one year. | |
| Chan et al. (2012) | Descriptive [Level 3A] | -Radiotherapist (*n*=1) -Radiation therapy system with 5 interfaces -Regular tasks during work tasks -Naturalistic | -Direct observation (30 hours) field of user tasks, workflow, interactions by 1 observer -Heuristic evaluation (HE) by 2 experts | -Usability problems -Device characteristics -Heuristic violations -Frequency and severity of violations | -Usability problems (*n*=75) into 14 categories: closures (*n*=2) to error (*n*=36) -Heuristic violations: Usability heuristics most commonly violated: error, consistency, memory -Severity ratings low (*n*=37), medium (*n*=37), | -Only 2 evaluators limits reliability and validity of the method -Unable to interact with device, limited validity of tasks -Simulated task after-hours -Tasks listed but not described or validated |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| | | setting, radio-therapy department after hours | using Zhang (2003) heuristic set | | high(*n*=18),<br>-Recommend HE as a viable aspect of procurement<br>-Notes low and med. severity increase cognitive load<br>-Recommend immediate interventions to mitigate safety of severe rated problems | -Radiotherapist not described<br>-Did not describe method for analysis of observation data.<br>-Inter-rater reliability not reported |
| Chiu, Vicente, Buffo-Sequeira, Hamilton, & McCrindle (2004) | Descriptive, non-experimental [Level 3 B] | -Pacemaker device programmers (*n*=42)<br>-Pacemaker programmer interfaces (*n*=7)<br>-Tasks-none<br>-Environment NA | -Self-administered survey: 20 Likert-type user perceptions of satisfaction<br>-Heuristic evaluation (HE) using Nielsen (1994) usability heuristics by 3 raters of 7 brands | -Satisfaction: User perception of ease of use, user satisfaction.<br>-Device characteristics and usability problems | -Survey identified significant differences between 7 programmers in user satisfaction, ease of programmer use, and component interface.<br>-Programmer interface does not meet user needs or adhere to usability principles.<br>-HE identified important safety, effectives, efficiency issues to inform manufactures of potential improvements | -Subjective nature of respondents and HE raters<br>-Small sample size for survey<br>-Only 3 evaluators limits reliability and validity of the HE method.<br>-Usability problems not reported but used to explain potential user satisfaction and ease of use perceptions. Used conceptual |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| | | | | | -HE provided context to explain usability problems from survey; device characteristics force "users cope with bad design by tailoring their activities and modifying their procedures" | model. |
| Etchells et al. (2006) | Descriptive [Level 3 B] | -Registered nurses (*n*=11)<br>-General infusion pumps (*n*=2)<br>-pump programming tasks<br>-Lab, high fidelity | -Developed and validated usability checklist and task check lists of programming pump<br>-Observations - nurses completing tasks with pump, recorded observations on usability checklist<br>-Secondary data analysis - Pump | -Usability problems<br>-Device characteristics<br>-Heuristic violations<br>-Frequency of violations | -Usability problems (*n*=5)<br>-Usability heuristics most commonly violated: error, consistency, memory<br>-Severity of usability problems low (*n*=2) and high (*n*=3)<br>-Identified usability problems with existing pump programming procedures<br>-Results used to modify training procedures and design pre-printed orders, and guide purchasing decisions | -Small sample limits generalizability<br>-Nurses changed patterns of behavior in response to being observed; observations may have led to mistakes/errors<br>-Use of high-fidelity simulation setting |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| | | | programming data downloaded and reviewed for patterns of error -Post usability interviews to gain insight into interactions | | | |
| Fairbanks, Bishop, Marks, & Shah (2007) | Experimental -Prospective crossover design [Level 1B] | -EMS providers (*n*=14) -Cardio-defibrillator (*n*=2) -4 tasks described -Lab | -Usability testing - Comparison of Medtronic LifePak10 and LifePak 12. Random assignment to first device. | -Effectiveness: task success, error rate -Satisfaction: Subjective user ratings of ease of task using questionnaire (open-ended user preferences, ratings of confidence, overall device rating) | -Task success (scale 0 failed to 4-excellent) LifePak 10 monitoring tasks had highest task success rate (mean=3.4); LifePak 10 cardioversion tasks had lowest task success rate (mean=1.6). High failure rate in synchronized cardio-version in one device -Error rate**:** experts observed incidence of undetected errors (*n*=5). Device did not communicate mode | -Use of simulated crime scene environment with Laerdal SimMan -Typical tasks were performed. -User previous experience was not controlled; all participates had experience with both models -Non-standardized measures used -Unable to re-produce stress, distractions, |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| | | | | | before second shock, contributing to undetected error. Continuous display of "sync" led user to erroneously believe mode was in synchronized<br>-Ease of use for LifePak 12 somewhat or very easy to use (71%); 85% (*n*=11) preferred LifePak 12 but the LifePak 10 was easier to learn. Lifepak 12 display visibility better than LifePak 10; Button configuration limited accurate use with gloved hands for mode selection, retrieving data, and printing results | and ambient environments of real-life; assumed real-life would present more stress and results in more errors<br>-Used only one mfg/brand, different versions (10 v. 12)<br>-No measures of efficiency |
| Garmer, Liljegren, Osvalder, & Dahlman (2002) | Experimental [Level 1 B] | -Nurses (*n*=18)<br>-PCA interface comparison (*n*=2)<br>-3 typical PCA interface tasks<br>-Lab | -Usability testing: Audio-video recorded 3 groups of nurses, completing tasks on PCA | -Effectiveness: error rate, mode error, undetected errors.<br>-Efficiency time to | -Error rate during task: Errors on existing interface were higher (*n*=28) compared to new interface (*n*=36); need to further improve the new interface. | -Users unable to think-aloud when they encounter problems, potentially workload considerations for future studies.<br>-Nature of the |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| | | | interfaces with "think aloud" protocol | complete tasks (seconds)<br>-Satisfaction: subjective ratings<br>-Other measures: frequency of manual use, need for help during test | -Mode error (data entered into wrong mode): Flow and volume to be infused errors existing interface (*n*=5) and new interface (*n*=7).<br>-Undetected errors (*n*=20) were ns between devices.<br>-Frequency of manual use more frequent with existing interface (*n*=29) than new (*n*=8).<br>-Need for help during test: Nurses who used the device infrequently, used the manual and needed help more frequently. Users gave up in group c, unable to complete tasks with help.<br>-Time to complete tasks (seconds): Existing interface (Median 260=seconds) compared to new interface (median=188 second) | information from the different user groups in usability tests can differ widely (experienced users vs. novice). |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| | | | | | (p<.05).<br>-Frequency of manual use and requests for help increased time to complete task.<br>-Satisfaction with interface: New interface was easier to learn and use; users felt it was difficult to change modes, how to understand symbols, and set flow rate. | |
| Ginsburg (2005) | Descriptive [Level 3 B] | -Nurses (*n*=14) and anesthetists (*n*=3)<br>-Smart™ infusion pumps (*n*=3) on 3 brands<br>-Common use tasks<br>-Lab | -Heuristic evaluation (1 expert rater) using 4 criteria sets<br>-Task analysis (nurses)<br>-Usability testing (*n*=17) | -Usability problems<br>-Device characteristics<br>-Heuristic violations, frequency of violation, unique usability problems (aggregated for each | -HFE and task identified strengths and weaknesses of each pump; Vendor A scored best<br>-Total error rates: Substantially fewer errors committed in Oncology, medical-surgical, and pediatrics with pump A.<br>-Total critical errors: Substantially fewer errors committed in Oncology, medical-surgical, and | -No novice users involved<br>-Small sample size tested in each clinical area<br>-Task and development well described, validated.<br>-Tasks broad to all clinical areas, not representative of specific clinical areas (i.e., focused on |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| | | | | device A-D)<br>-Effectiveness: total error rate, total critical errors, total critical undetected errors<br>-Satisfaction: User preference | pediatrics with pump A.<br>-Total critical undetected errors: Fewer errors undetected pediatrics with pump A.<br>-Undetected critical errors and usability errors committed most frequently with Pump C.<br>-User preference highest for Pump A.<br>-Use of complementary heuristic evaluations and usability testing can guide design change to improve usability but can effectively be used by organizations during medical device procurement decisions.<br>-Organizations should tailor training program to inform users of usability problems. | common programming)<br>-Heuristic, usability, errors, and user preference findings were aggregated and summarized in tables, not specifically described.<br>-User testing environment varied.<br>-Pump order was not counter-balanced across participants in each area<br>-Observed errors recorded by hand as observed; errors may have been missed since video recording was not incorporated)<br>-Only 1 evaluator limits reliability and validity of the HFE method<br>-Results were used to |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| | | | | | | make large-scale purchase decision |
| Graham et al. (2004) | Non-research [Level 5 A] | -User- NA<br>-Device- 3-channel infusion pump (*n*=1)<br>-Tasks-none<br>-Environment – NA | -Heuristic evaluation (HE) by 4-raters on one –channel pump using heuristic set by Zhang (2003)* | -Heuristic violations<br>-Frequency of violations<br>-Severity ratings<br>-Usability problems<br>-Device characteristics | -Heuristic violations (*n*=231)<br>-Severity ratings: catastrophic (*n*=9), major (*n*=61), minor (*n*=48), cosmetic (*n*=11)<br>-Inter-rater reliability kappa test (range .52-.62, mean=.60, *p*=.01)<br>-Interface source of *Primary screen* (62 violations, 42%) *Options screen* (25 violations, 17%)<br>-Usability heuristics most commonly violated: consistency, language, error, match<br>-End users must be vigilant about the potential for making errors across numerous interface aspects of device | -Evaluators included 2 HFE and 1 veteran ICU RN<br>-Moderate kappa results<br>-Evaluation considered only ICU nurse users in light of proportion of users to experience a problem, impact, persistence, & severity |
| Lin et al., | Experimental | -Nursing | -Cognitive task | -Efficiency: | -Evaluations identified | -Issues relating to |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| (1998) | [Level 1C] | students ($n$=12) -Abbott Lifecare 4100 PCA ($n$=2) -PCA Programming tasks -Lab and naturalistic environment (hospital) | analysis -Bench tests -Field observations -Re-engineered new user interface then tested -Users completed 12 tasks on each interface (old, new) | mean programming time (minutes) -Efficiency: subjective workload (NASA-TLX) -Effectiveness: programming errors -user preference of device interface. | new design requirements. Old and new were compared in empirical study. -New interface results showed significantly fewer errors ($x^2(1) = 3.33$, $p < 0.05$), faster programming times ($F(1,11) = 6.85$, $P < 0.025$), and lower mental workload ratings ($x^2 (1) = 4.45$, $p < 0.025$). -All 12 users expressed strong preference for new interface. | transfer from one device to the other were not controlled -Study included only novices, limiting generalizability -Test of a new prototype in a simulated setting -Device was redesigned by manufacturer. Device tested may not compare to new device, limiting generalizability to only the proto-type device. |
| Lin, Vicente, & Doyle (2001) | Experimental -Mixed design 2x2x3x2 [Level 1 B] | -Registered nurses ($n$=12) -PCA device and interface -PCA Programming tasks -Lab | -Usability testing - Commercial PCA (Old) compared to prototype of new interface design (New). | -Efficiency: task completion time (minutes), task completion (percent). | -Task completion time: New interface was statistically faster than old for task completion time ($F(1,10)=12.17$, $p$=.006) . -Mental effort: NASA-TLX lower for the second repetition than the first | -Single specialty (recovery room) of nurses with frequent programming experience selected, results not generalizable to other populations. |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| | | | Order of interface and mode counterbalanced. | -Efficiency: mental effort (NASA-TLX)<br>-Effectiveness: errors<br>-User preference of device interface | ($F$)1,10)=3.27, p=0.03). Workload reduced significant with repetition in the new interface ($F$(2, 20)=8.62, p=0.002).<br>-Errors: Old interface was involved in statistically more errors ($n$=29, $p$<.05) than new interface ($n$=13). New interface was not involved in drug concentration errors.<br>-User interface preference: nurses ($n$=9) preferred new interface compared to Old interface ($n$=1), No preference ($n$=2). | -Simulation did not control for interruptions which are common in natural setting.<br>-Device data input devices were not identical<br>-Study included only one device.<br>-NASA-TLX not described. |
| Liu, Tech, & Osvalder (2004) | Experimental [Level 1B] | -Registered nurses ($n$=6)<br>-Ventilator interface ($n$=2)<br>-Task-detection and interpretation of display data<br>-Lab | -Experimental usability test<br>-Comparison and evaluation of numerical ventilator display to prototype graphical user | -Effectiveness: error rates, expert subjective severity of deviations<br>-Efficiency: detection time<br>-Satisfaction: | -Error rates for interpreting deviations and assessing the overall situation, were not significant between numerical and GUI. Error rates for interpreting the meaning of deviation improved with the GUI ($p$<.05) compared to | -Only one ventilator interface and one prototype on one mode (volume control) were included in the study.<br>-Use of simulated test environment may not reflect the distracted |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| | | | interface (GUI) design<br>-Interviews - expert nurse users. Sequence of six tasks randomized. | display type reference | numerical display.<br>-Severity of deviations was easier to detect using the HUI and the majority preferred the GUI.<br>-Detection time between use of numerical and GUI *n.s.* ($t \geq 1.14$, $df=19$, $p>0.5$). GUI better at helping to interpret meaning of parameter deviations and does not contribute to detection time.<br>-Majority of users preferred graphical display. | work in ICU or real world work environment. |
| Nemeth, Nunnally, Bitan, Nunally, & Cook (2009) | Non-experimental [Level 3 A] | -Registered nurses, experienced ($n=19$)<br>-General infusion pumps ($n=4$)<br>-Programming tasks for set- | -Expert usability review<br>-Experimental user testing - video recorded tasks with think aloud. Pump order counterbalance | -User characteristics<br>-Device characteristics<br>-User perceptions of device | -Observations categorized into 4 themes:<br>-Programming by users showed no correlation between clinical experience and ability to program any of the pumps under consideration.<br>-Field observations | -User experience was not observed to improve use<br>-Many use patterns in practice may not be generalized to other areas or practices. |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| | | up, administration, and maintenance of fluids -Lab and naturalistic setting | d -Observations in naturalistic setting and secondary analysis of pump programming and use data -Analysis of recent adverse event reports in the MAUDE database | | reflected diverse use patterns across services that required *ease of use pumps did not offer*. -Upon review of a final candidate pump, purchasing preferences superceded clinical considerations. -Study results were used to inform institutional decision-making selection of commercially available infusion device. | |
| Nunnally & Bitan (2006) | Non-Experimental [Level 3 B] | -Registered nurses, experienced (*n*=19) -General infusion pumps (*n*=4) from 4 manufacturers -Programming and set up for automated secondary | -Experimental user testing - audio-video recorded tasks with think aloud then secondary data analysis of pump operational log data -Analysis of | -Effectiveness: task completion rate -Efficiency: interface pattern tracking -Satisfaction: display type reference | -User testing: Substantial task completion fail rate. User behavior incongruent with device-user cognition observed suggesting confusion, use of trial and errors, which may limit efficiency. -Subjects failed to accurately complete tasks (53%) of scenarios; subjects became lost in | -Task data set was insufficient in size for between pump comparisons, tasks not validated. -Tasks, use environment not described. -Use of simulated lab setting limit -Participant behaviors may have been |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| | | infusions -Lab | medical device adverse event data 2003-2004 specific infusion device | | programming pathways (29%). -Pump operation logs revealed 124 secondary infusions in 26 different pumps; pumps switch from secondary to primary in 85% of events. -*User events*: MAUDE data revealed 137 secondary infusion cases but were limited in detail; calculation of prevalence not possible. MAUDE reports are poorly detailed and lack detail to inform analyses. | affected by observation during tasks; subjects had little or no familiarity with devices -MAUDE data set and pump operation logs were limited |
| Obradovich & Woods (1996) | Descriptive, qualitative [Level 3 B] | -Registered nurses, home health (*n* - not reported) -Syringe infusion pump -Typical infusion programming | -In-depth nurse interviews -Device bench tests device exploring device behavior, displays, tasks, contexts of use | -Device characteristics (limitations) that create or enhance error -Error-prone tasks -Tailoring strategies to | -HCI deficiencies are device characteristics that produce or augment error potential such as classic HCI deficiencies: limitations in user-device feedback and behavior, ambiguous alarms, complex or arbitrary | Note: Author did not report study limitations |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| | | and teaching tasks -Lab and naturalistic (home care setting) | -Observations of nurses' device programming | compensate device limitations | operational sequencing, and context-based operating modes. HCI deficiencies lead to getting lost in complex command sequences. -Tailoring strategies used to work-around error prone tasks and device deficiencies: nurses developed a user manual and checklist for patients and changes, modified, and introduced new procedures for patient-users. -Latent errors can result from poor usability. | |
| Rogers, Mykitshyn, Campbell, & Fisk (2001) | Non-research [Level 5 A] | -Users-NA -Blood glucose meter (*n*=1) -Typical tasks (*n*=3) -Environment-NA | -Task analysis -Instructional analysis -Expert usability evaluation with typical users and satisfaction report | -Task sequence, number of tasks -Readability of user manual and instructional | -Task analysis detailed 52 task steps, user task/knowledge, feedback provided to user, and potential problems -Instructional analyses: --user manual scored 8th grade level, readable by | -Note: Author did not report any study limitations -Study limited to one device -Method for usability evaluation and measures for usability |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| | | | | video using Flesch-Kincaid Grade Level Analysis (global measure of readability)<br>-User satisfaction | 58% of US population<br>--lanceted instructions scored 6th grade, readable by 72%<br>--test strip instructions scored 9th grade readable by 51%.<br>--Instructional video switches back and forth, aspects of procedure may be missed. Older adults had more difficulty in completing tasks<br>-Satisfaction reports: 70% of report related to problems using system; users average 2.5 brands where dissatisfaction led to try new meters; 50% used instructional manual as primary source of instruction.<br>-Findings were used to make recommendations for system design (modify strips, meter, features, | satisfaction not described |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| | | | | | blood sampling procedure, and major systems) and Instructional design (readability, vocabulary, video, redundancy). | |
| Trbovich, Pinkney, Cafazzo, & Easty (2010) | Experimental -3x7 design, repeated measures [Level 1B] | -Nurses (*n*=24)<br>-General infusion pumps, 4 brands<br>-Infusion programming tasks (4 tasks/nurse)<br>-Lab, high fidelity | -Nurses delivered infusions with each of 3 pumps in high-fidelity in-patient simulation lab; tasks counterbalanced with pump type. Each nurse completed 21 infusions. | -Effectiveness: programming accuracy, secondary infusion error, error resolution | -Programming accuracy for continuous infusions: 203 of 216 infusions (94%) were accurate. NS difference across pump types.<br>-Programming accuracy, intermittent infusions: Tasks were significantly more accurate ($p<.01$) with smart pump and barcode pump.<br>-Secondary infusion error: Error rates (mean=55.6%) were high across all pumps. Error rate *ns*. with pump type.<br>-Error resolution: Users remedied planted drug errors on 43 of 72 entries | -Small sample size and use of simulated environment limit generalizability<br>-Planted errors may have affected nurse behavior<br>-Nurses has no previous experience with the technology limiting generalizability to novice users<br>-Tasks validated prior to use.<br>-No measure of efficiency or satisfaction included. |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| | | | | | (60%). Pump type did not significantly impact wrong dose errors. | |
| Turley, Johnson, Smith, Zhang, & Brixey (2006) | Non-research [Level 5 A] | -User-NA<br>-Device- infusion pump operating manual (*n*=5)<br>-Tasks-none<br>-Environment – NA | -Heuristic evaluation (HE) with 2 raters using Zhang (2003) heuristic set | -Heuristic violations<br>-Frequency of violations<br>-Severity ratings<br>-Usability problems<br>-Device characteristics | -Heuristic violations (range 7-36) Most frequently violated were minimize memory load (*n*=54) and prevent errors (*n*=36)<br>-Severity ratings highest for pump E; Pump E had the most major and catastrophic severity ratings (62.5%)<br>-Pump C received highest recommendations, had fewest heuristic violations and least severe ratings<br>-Findings used to support pre-purchasing decision-making<br>-Simple and cost effective method | -Only 2 evaluators (1 for heuristics, 2 for severity ratings) limits reliability and validity of the HE method and is inconsistent with literature<br>-Application of HE as a proxy measure of usability to operating manuals needs to be validated<br>-HE was depended on quality of content provided by manufacturer within the operating manual |
| Zhang, Johnson, Patel, | Non-research [Level 5 A] | -User-none<br>-Infusion pumps (*n*=2) | -Heuristic Evaluation (HE) by 4 raters | -Heuristic violations<br>-Frequency of | -Pump 1 Heuristic violations (*n*=192), usability problems | Note: Author did not report study limitations |

| Source | Design/ Approach [Quality of evidence rating]* | Formation of Usability: -User (Sample) -Device -Task/Activity -Environment | Method/ Intervention | Usability attributes/ Variable measures | Key Findings | Limitations |
|---|---|---|---|---|---|---|
| Paige, & Kubose (2003) | | -Tasks-none<br>-Environment – not described | and comparison of 2 pump brands using Zhang (2003) heuristic set | violations<br>-Severity ratings<br>-Usability problems<br>-Device characteristics | (*n*=89); catastrophic usability problems (*n*=2). Most frequently violated accounted for 64% of violations: consistency (*n*=53), visibility (*n*=28), feedback (*n*=22), match (*n*=21).<br>-Pump 2 Heuristic violations (*n*=121), usability problems (*n*=52), catastrophic usability problems (*n*=1). Most frequently violated accounted for 54% of violations: visibility (*n*=29), memory (*n*=19), consistency (*n*=17).<br>-Severity of usability problems greater in Pump 2.<br>-Both pumps had major and minor usability problems. | |

* Quality of evidence rating. This bracketed information lists the strength of evidence rating (Newhouse, Dearholt, Poe, Pugh, & White, 2007) assigned to each design. Strength of evidence level 1 to 5 and quality of evidence high-A, god-B, or low-C.

**APPENDIX B**
**INFORMED CONSENT**

University of Central Florida

***Impact of Interruption Frequency on Nurses' Performance, Cognition, and Satisfaction***

**Informed Consent**

| | |
|---|---|
| Principal Investigator(s): | Kristi R. Campoe, MSN, RN |
| Sub-Investigator(s): | Robert Wagner, AA-R<br>Karen Giuliano, PhD, RN |
| Faculty Supervisor: | Steven Talbert, PhD, RN |
| Sponsor: | Nova Southeastern University |
| Investigational Site(s): | Nova Southeastern University<br>3632 Queen Palm Drive<br>Tampa, FL 33619 |

**Introduction:** Researchers at the University of Central Florida (UCF) study many topics. To do this we need the help of people who agree to take part in a research study. You are being invited to take part in a research study which will include about 9 people in Florida. You have been asked to take part in this research study because you are a registered nurse. You must be 18 years of age or older to be included in the research study.

**Kristi R. Campoe, MSN, RN**: The person doing this research is a College of Nursing doctoral candidate from the University of Central Florida. Because the researcher is a graduate she is being guided by Dr. Steven Talbert, a UCF faculty supervisor in the College of Nursing.

**What you should know about a research study:**

- Someone will explain this research study to you.
- A research study is something you volunteer for.
- Whether or not you take part is up to you.
- You should take part in this study only because you want to.
- You can choose not to take part in the research study.
- You can agree to take part now and later change your mind.
- Whatever you decide it will not be held against you.
- Feel free to ask all the questions you want before you decide.



UCF University of Central Florida IRB
IRB NUMBER: SBE-14-10205
IRB APPROVAL DATE: 4/2/2014
IRB EXPIRATION DATE: 4/1/2015

**Purpose of the research study:** The purpose of this research study is to determine the impact of interruption frequency on nurses' performance, cognition, and satisfaction and to determine nurses' perception of the impact of interruption frequency after interactions with patient-controlled analgesia devices. Nurse are interrupted frequently during medication administration. This study will attempt to improve our understanding of the impact of frequent interruption during medication administration.

**What you will be asked to do in the study:** While sitting in a private room, you will complete a brief survey that asks you information about yourself, your workplace, and your experience with patient-controlled analgesia systems. You will also receive training on the patient-controlled analgesia system that will be used in this study. In the simulation laboratory, you will be given several set of tasks by the principal investigator to complete while programming a patient-controlled analgesia system. You may be interrupted while performing the assigned tasks. After you complete sets of tasks, you will complete another brief survey about your interaction with the patient controlled analgesia where you will describe your performance, satisfaction, and cognitive workload. This process will be repeated four times in total. Your responsibility is to complete the tasks involving the patient controlled analgesia system and respond to interruptions as you would in a real world setting. Your interactions with the patient controlled analgesia system will be audio and video recorded and documented by the principal investigator. After all tasks are finished, there will be a brief follow up interview conducted by the principal investigator after you have completed all of the tasks. The interview will be audio recorded. Once the interview is completed, the experiment is completed and your responsibility is fulfilled.

**Location:** The study will take place at Nova Southeastern University Tampa Campus in the Anesthesiology Assistant laboratory.

**Time required:** We expect that you will be in this research study for approximately one hour and a half on one day, during a single session. You will be offered a break during the time you are here.

**Audio or video taping:** You will be audio taped during this study. If you do not want to be audio taped, you will not be able to be in the study. Discuss this with the researcher or a research team member. If you are audio taped, the tape will be kept in a locked, safe place. The tape will be erased or destroyed after the recording is analyzed and the study is complete.

You will be videotaped during this study. If you do not want to be videotaped, you will not be able to be in the study. Discuss this with the researcher or a research team member. If you are video taped, the tape will be kept in a locked, safe place. The tape will be erased or destroyed after the recording is analyzed and the study is complete.

**Funding for this study:** This research study is being paid for by Nova Southeastern University.

**Risks:** There are no reasonably foreseeable risks or discomforts involved in taking part in this study. There is a small risk that people who take part will develop what is ordinarily referred to as simulator sickness. It occurs once in awhile to people who are exposed to prolonged continuous testing in simulated environments. Symptoms consist of nausea and a feeling of being light- headed. The risk is minimized as a result of the short duration of each session in the simulator. If you experience any of the symptoms mentioned, please tell the researcher and remain seated until the symptoms disappear.

**Benefits:** There are no expected benefits to you for taking part in this study.

**Compensation or payment:** There is no cost to you to participate in this research study. Participants may expect to spend between 60 and 90 minutes performing experimental tasks. Compensation for your participation will be $45 retail gift card after completing the entire experiment and interview. If you do not complete the entire experiment, you will be compensated for time spent on study activities: $15 per 30 minutes for a maximum compensation for participation is one $45 retail gift card for completing the study.

**Confidentiality:** We will limit your personal data collected in this study to people who have a need to review this information. We cannot promise complete secrecy. Organizations that may inspect and copy your information include the IRB and other representatives of UCF.

**Right to withdraw from the study**: You have the right to withdraw from the study at any time without penalty. Notify the principal investigator anytime during the study if you decide to no longer participate.

**Study contact for questions about the study or to report a problem:** If you have questions, concerns, or complaints, or think the research has hurt you, talk to Kristi R. Campoe, Graduate Student, College of Nursing (407) 823-2744 or Dr. Steven Talbert, Faculty Supervisor, College of Nursing at (407) 823-2744 or by email at steven.talbert@ucf.edu.

**IRB contact about your rights in the study or to report a complaint:** Research at the University of Central Florida involving human participants is carried out under the oversight of the Institutional Review Board (UCF IRB). This research has been reviewed and approved by the IRB. For information about the rights of people who take part in research, please contact: Institutional Review Board, University of Central Florida, Office of Research & Commercialization, 12201 Research Parkway, Suite 501, Orlando, FL 32826-3246 or by telephone at (407) 823-2901. You may also talk to them for any of the following:

- Your questions, concerns, or complaints are not being answered by the research team.
- You cannot reach the research team.
- You want to talk to someone besides the research team.
- You want to get information or provide input about this research.

**Results of the research:** To receive a copy of the study results, contact the principal investigator and a summary of the results will be emailed to you.

**APPENDIX C**
**SIMULATION CASE SCENARIOS WITH INTERRUPTED TASKS**

**Scenarios**

Scenarios have been developed utilizing existing scenarios with the Medical Education Technologies inc., (METI) Learning system® high-fidelity patient simulator. Each scenario simulates medical and surgical patients commonly receiving opioids via patience-controlled analgesia (PCA) in the general adult medical-surgical patient care setting. Each scenario involves the participant receiving a patient care report, setting up the PCA device according to physician orders using standardized order sets, and making changes to PCA programming with new orders or after changes in patient status.

Scenarios will be described to each participant following a standardized communication tool known as SBAR in which the nurse will received information on the patients' current situation, medical background, current assessment, and current orders with recommendations. PCA orders for patients will be included with a written report and physician orders or standardized order set. All scenarios and PCA orders (order sets) will be reviewed by two content experts prior to the study. Scenarios with tasks and interruptions will be preliminarily tested before being piloted with study participants.

1. Scenario one involves a 48-year old male who is recently admitted with acute pancreatitis with orders to receive morphine sulfate via PCA.
2. Scenario two involves a 56-year old male who is postoperative partial gastrectomy with orders to receive Hydromorphone via PCA.
3. Scenario three involves a 28-year old female who was admitted in sickle cell crisis with orders to receive Morphine sulfate via PCA.
4. Scenario four involves a 61-year old female who is postoperative open reduction internal fixation (ORIF) right hip with orders to receive Hydromorphone via PCA.

**PCA Tasks**

PCA tasks involve main functions of the PCA infusion pump for medication administration and monitoring commonly used in physician standardized PCA order sets. Each task was developed according to the Baxter PCA II Pump Operator's Manual (Baxter Healthcare, 1993).

Task 1. Initial pump set up with initial PCA only mode programming, administration of bolus dose, verify Rx, and start infusion.

Subtask 1a. Pump set up with continuous mode programming

Subtask 1b. Initiation of loading (bolus) dose

Subtask 1c. Verify RX

Subtask 1d. Initiation of infusion

Task 2. Change PCA orders: Pump set up and Basal/PCA mode programming, administration of bolus dose, verify Rx, and start infusion.

Subtask 2a. Pump set up with continuous mode programming

Subtask 2b. Initiation of loading (bolus) dose

Subtask 2c. Verify RX

Subtask 2d. Initiation of infusion

Task 3. Change PCA orders: Pump set up with continuous initial programming, administration of bolus dose, verify Rx, and start infusion.

Subtask 3a. Pump set up with continuous mode programming

Subtask 3b. Initiation of loading (bolus) dose

Subtask 3c. Verify RX

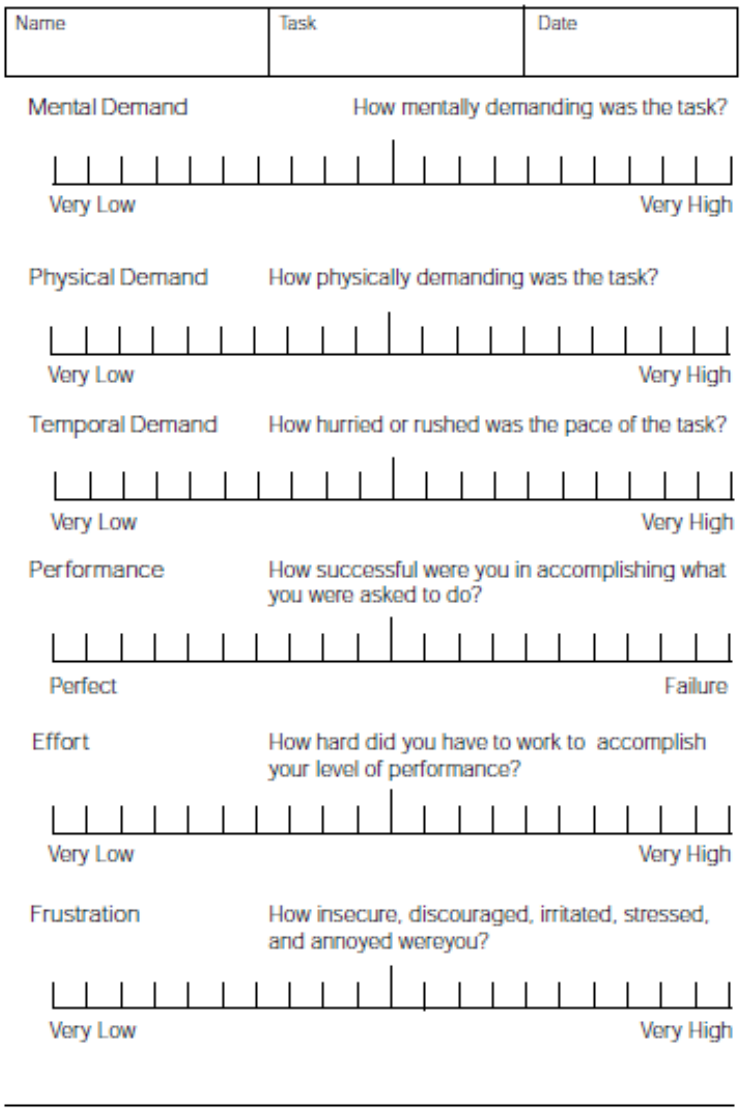
Subtask 3d. Initiation of infusion

**APPENDIX D**
**NASA TASK LOAD INDEX**

**Figure 8.6**

## *NASA Task Load Index*

*Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.*



**APPENDIX E**
**SEMI-STRUCTURED INTERVIEW GUIDE**

**Instructions to participant: The following three questions require that you rate the impact of interruption from zero to four.**

You were interrupted at varying rate, from no interruptions up to 6 interruptions per 10 minutes during your interactions with the PCA. How would you rate the *overall impact* of interruption frequency on the following during your interactions with the PCA on your:

1. On your performance?

   No impact-(0) Low impact-(1) Moderate impact-(3) High impact-(4)

2. On your satisfaction?

   No impact-(0) Low impact-(1) Moderate impact-(3) High impact-(4)

3. On your subjective workload (effort, frustration, attention, perception, memory levels)?

   No impact-(0) Low impact-(1) Moderate impact-(3) High impact-(4)

**Follow up questions, regarding frequency of interruption:**

4. Describe the frequency of interruptions in your current work environment.
5. Describe how the frequency of interruptions you experienced today compares to your current work environment.
6. Using your own words, how would you describe the *impact of frequency of interruptions* during your interactions with the PCA. At work? Today?
7. Describe how frequency of interruption created physical, temporal or other demands on you.

**Regarding auditory and visual interruptions:**

8. Describe auditory and visual interruptions in your current work environment.
9. Describe any auditory and visual interruptions you experienced today.

10. Describe how the auditory and visual interruptions you experienced today compares to your current work environment.
11. Are there additional comments you would like to add regarding your participation in the experiment or with regard to the interview?

**Interview Closure**:

1. The interview is complete. Thank the participant.
2. Complete the interview by requesting permission to follow up by phone call for additional questions that may arise. ___YES ___NO

   Phone Number _____
3. Provide the incentive as approved and have the participant sign incentive receipt.

**APPENDIX F**
**NSU IRB APPROVAL**

NOVA SOUTHEASTERN UNIVERSITY

# MEMORANDUM

To: Kristi Campoe, MSN, RN, CMSRN, CPHQ
HPD – College of Nursing

From: David Thomas, M.D., J.D.
Chair, Institutional Review Board

Date: May 22, 2014

Re: *Impact of Interruption Frequency on Nurse Performance, Satisfaction, and Cognition* – NSU IRB No. 05021436Exp.

---

I have reviewed the revisions to the above-referenced research protocol by an expedited procedure. On behalf of the Institutional Review Board of Nova Southeastern University, *Impact of Interruption Frequency on Nurse Performance, Satisfaction, and Cognition* is approved in keeping with expedited review category # 6 and #7. Your study is approved on **May 22, 2014** and is approved until **May 21, 2015**. You are required to submit for continuing review by **April 21, 2015**. As principal investigator, you must adhere to the following requirements:

1) CONSENT: You must use the stamped (dated consent forms) attached when consenting subjects. The consent forms must indicate the approval and its date. The forms must be administered in such a manner that they are clearly understood by the subjects. The subjects must be given a copy of the signed consent document, and a copy must be placed with the subjects' confidential chart/file.

2) ADVERSE EVENTS/UNANTICIPATED PROBLEMS: The principal investigator is required to notify the IRB chair of any adverse reactions that may develop as a result of this study. Approval may be withdrawn if the problem is serious.

3) AMENDMENTS: Any changes in the study (e.g., procedures, consent forms, investigators, etc.) must be approved by the IRB prior to implementation.

4) CONTINUING REVIEWS: A continuing review (progress report) must be submitted by the continuing review date noted above. Please see the IRB web site for continuing review information.

5) FINAL REPORT: You are required to notify the IRB Office within 30 days of the conclusion of the research that the study has ended via the IRB Closing Report form.

The NSU IRB is in compliance with the requirements for the protection of human subjects prescribed in Part 46 of Title 45 of the Code of Federal Regulations (45 CFR 46) revised June 18, 1991.

Cc: Dr. Robin Chard
Ms. Jennifer Dillon

Institutional Review Board
3301 College Avenue • Fort Lauderdale, Florida 33314-7796
(954) 262-5369 • Fax: (954) 262-3977 • Email: *irb@nsu.nova.edu* • Web site: *www.nova.edu/irb*

University of Central Florida

# *Impact of Interruption Frequency on Nurses' Performance, Cognition, and Satisfaction*

## Informed Consent

| | |
|---|---|
| Principal Investigator(s): | Kristi R. Campoe, MSN, RN |
| Sub-Investigator(s): | Robert Wagner, AA-R<br>Karen Giuliano, PhD, RN |
| Faculty Supervisor: | Steven Talbert, PhD, RN |
| Sponsor: | Nova Southeastern University |
| Investigational Site(s): | Nova Southeastern University<br>3632 Queen Palm Drive<br>Tampa, FL 33619 |

**Introduction:** Researchers at the University of Central Florida (UCF) study many topics. To do this we need the help of people who agree to take part in a research study. You are being invited to take part in a research study which will include about 9 people in Florida. You have been asked to take part in this research study because you are a registered nurse. You must be 18 years of age or older to be included in the research study.

**Kristi R. Campoe, MSN, RN**: The person doing this research is a College of Nursing doctoral candidate from the University of Central Florida. Because the researcher is a graduate she is being guided by Dr. Steven Talbert, a UCF faculty supervisor in the College of Nursing.

**What you should know about a research study:**

- Someone will explain this research study to you.
- A research study is something you volunteer for.
- Whether or not you take part is up to you.
- You should take part in this study only because you want to.
- You can choose not to take part in the research study.
- You can agree to take part now and later change your mind.
- Whatever you decide it will not be held against you.
- Feel free to ask all the questions you want before you decide.

**Purpose of the research study:** The purpose of this research study is to determine the impact of interruption frequency on nurses' performance, cognition, and satisfaction and to determine nurses' perception of the impact of interruption frequency after interactions with patient-controlled analgesia devices. Nurse are interrupted frequently during medication administration. This study will attempt to improve our understanding of the impact of frequent interruption during medication administration.

**What you will be asked to do in the study:** While sitting in a private room, you will complete a brief survey that asks you information about yourself, your workplace, and your experience with patient-controlled analgesia systems. You will also receive training on the patient-controlled analgesia system that will be used in this study. In the simulation laboratory, you will be given several set of tasks by the principal investigator to complete while programming a patient-controlled analgesia system. You may be interrupted while performing the assigned tasks. After you complete sets of tasks, you will complete another brief survey about your interaction with the patient controlled analgesia where you will describe your performance, satisfaction, and cognitive workload. This process will be repeated four times in total. Your responsibility is to complete the tasks involving the patient controlled analgesia system and respond to interruptions as you would in a real world setting. Your interactions with the patient controlled analgesia system will be audio and video recorded and documented by the principal investigator. After all tasks are finished, there will be a brief follow up interview conducted by the principal investigator after you have completed all of the tasks. The interview will be audio recorded. Once the interview is completed, the experiment is completed and your responsibility is fulfilled.

**Location:** The study will take place at Nova Southeastern University Tampa Campus in the Anesthesiology Assistant laboratory.

**Time required:** We expect that you will be in this research study for approximately one hour and a half on one day, during a single session. You will be offered a break during the time you are here.

**Audio or video taping:** You will be audio taped during this study. If you do not want to be audio taped, you will not be able to be in the study. Discuss this with the researcher or a research team member. If you are audio taped, the tape will be kept in a locked, safe place. The tape will be erased or destroyed after the recording is analyzed and the study is complete.

You will be videotaped during this study. If you do not want to be videotaped, you will not be able to be in the study. Discuss this with the researcher or a research team member. If you are video taped, the tape will be kept in a locked, safe place. The tape will be erased or destroyed after the recording is analyzed and the study is complete.

**Funding for this study:** This research study is being paid for by Nova Southeastern University.

**Risks:** There are no reasonably foreseeable risks or discomforts involved in taking part in this study. There is a small risk that people who take part will develop what is ordinarily referred to as simulator sickness. It occurs once in awhile to people who are exposed to prolonged continuous testing in simulated environments. Symptoms consist of nausea and a feeling of being light- headed. The risk is minimized as a result of the short duration of each session in the simulator. If you experience any of the symptoms mentioned, please tell the researcher and remain seated until the symptoms disappear.

**Benefits:** There are no expected benefits to you for taking part in this study.

**Compensation or payment:** There is no cost to you to participate in this research study. Participants may expect to spend between 60 and 90 minutes performing experimental tasks. Compensation for your participation will be $45 retail gift card after completing the entire experiment and interview. If you do not complete the entire experiment, you will be compensated for time spent on study activities: $15 per 30 minutes for a maximum compensation for participation is one $45 retail gift card for completing the study.

**Confidentiality:** We will limit your personal data collected in this study to people who have a need to review this information. We cannot promise complete secrecy. Organizations that may inspect and copy your information include the IRB and other representatives of UCF.

**Right to withdraw from the study**: You have the right to withdraw from the study at any time without penalty. Notify the principal investigator anytime during the study if you decide to no longer participate.

**Study contact for questions about the study or to report a problem:** If you have questions, concerns, or complaints, or think the research has hurt you, talk to Kristi R. Campoe, Graduate Student, College of Nursing (407) 823-2744 or Dr. Steven Talbert, Faculty Supervisor, College of Nursing at (407) 823-2744 or by email at steven.talbert@ucf.edu.

**IRB contact about your rights in the study or to report a complaint:** Research at the University of Central Florida involving human participants is carried out under the oversight of the Institutional Review Board (UCF IRB). This research has been reviewed and approved by the IRB. For information about the rights of people who take part in research, please contact: Institutional Review Board, University of Central Florida, Office of Research & Commercialization, 12201 Research Parkway, Suite 501, Orlando, FL 32826-3246 or by telephone at (407) 823-2901. You may also talk to them for any of the following:

- Your questions, concerns, or complaints are not being answered by the research team.
- You cannot reach the research team.
- You want to talk to someone besides the research team.
- You want to get information or provide input about this research.

**Results of the research:** To receive a copy of the study results, contact the principal investigator and a summary of the results will be emailed to you.

Email to Nurse Administrators and Educators

Dear Nurse Leader:

My name is Kristi Campoe. I work at Nova Southeastern University (NSU) College of Nursing as an Assistant Professor in the RN to BSN/MSN program. I also attend the University of Central Florida (UCF) where I am pursuing my PhD in Nursing.

I have completed all of the course work required for this degree and am beginning my dissertation. My dissertation research involves studying the effects of frequency of interruption on medical-surgical registered nurses' performance, satisfaction, and cognition during administration of patient-controlled analgesia (PCA) in a stimulated setting. Nurses will complete a series of PCA programming tasks in a high-fidelity simulation environment. Their performance will be video and audio recorded and then each nurses will complete a questionnaires and a brief interview. The results of the study aim to improve our understanding of nurses' interactions with PCA in the complex medical-surgical work environment. Recruitment for this study will not begin until the study is approved by the IRB at NSU and at UCF. Also, this study was funded by a grant from Nova Southeastern University. Therefore, nurses who participate will receive a retail gift card for participating.

I need your assistance in recruiting qualified medical-surgical registered nurses for this study. I am requesting that you share this request and recruitment flyer with medical-surgical unit leaders and nurses. The recruitment flyer can be emailed to nurses and posted in common areas. Registered nurses must:

(a) be employed 24 or more hours per week on average in a medical-surgical unit;
(b) have at least 6 months experience in adult medical-surgical nursing; *and*
(c) indicate/self-report patient-controlled analgesia device/system programming use at least four times per month on their medical-surgical unit.

It is important the medical-surgical registered nurses interested in participating contact me directly and soon as the study begins in a few weeks! The results of this study will be presented as Nova Southeastern University in an Open House format after the study is completed. You and your nursing team will be invited to attend this event.

I thank you in advance for supporting recruitment of this funded dissertation study.

Sincerely,

Kristi R. Campoe, MSN, RN, CMSRN, CPHQ
Doctoral Candidate
University of Central Florida
kcampoe@knights.ucf.edu

Dr. Steven Talbert, PhD, RN
Assistant Professor
University of Central Florida
talbert@knights.ucf.edu

**APPENDIX G**
**UCF IRB APPROVAL**

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

# Approval of Human Research

From: **UCF Institutional Review Board #1**
**FWA00000351, IRB00001138**

To: **Kristi Campoe**

Date: **April 02, 2014**

Dear Researcher:

On 4/2/2014, the IRB approved the following human participant research until 4/1/2015 inclusive:

| | |
|---|---|
| Type of Review: | UCF Initial Review Submission Form |
| Project Title: | The Impact of Interruption Frequency on Nurses' Performance, Cognition, and Satisfaction |
| Investigator: | Kristi Campoe |
| IRB Number: | SBE-14-10205 |
| Funding Agency: | Nova Southeastern University |
| Grant Title: | Impact of Interruption Frequency on Nurse Performance, Satisfaction, and Cognition |
| Research ID: | N/A |

The scientific merit of the research was considered during the IRB review. The Continuing Review Application must be submitted 30days prior to the expiration date for studies that were previously expedited, and 60 days prior to the expiration date for research that was previously reviewed at a convened meeting. Do not make changes to the study (i.e., protocol, methodology, consent form, personnel, site, etc.) before obtaining IRB approval. A Modification Form **cannot** be used to extend the approval period of a study. All forms may be completed and submitted online at https://iris.research.ucf.edu .

If continuing review approval is not granted before the expiration date of 4/1/2015, approval of this research expires on that date. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

Use of the approved, stamped consent document(s) is required. The new form supersedes all previous versions, which are now invalid for further use. Only approved investigators (or other approved key study personnel) may solicit consent for research participation. Participants or their representatives must receive a copy of the consent form(s).

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

On behalf of Sophia Dziegielewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

Joanne Muratori

IRB Coordinator

**APPENDIX H**
**NASA-TLX PERMISSION**

RE: Permission to use and access computerized NASA-TLX

DELETE REPLY REPLY ALL FORWARD

mark as unread

Chrestenson, Kim L. (ARC-VP)[DELTHA-CRITIQUE] <kim.l.chrestenson@nasa.gov>
Tue 9/3/2013 1:52 PM

To: Kristi Campoe <kcampoe@knights.ucf.edu>;
Cc: Gore, Brian F. (ARC-TH)[SAN JOSE STATE UNIVERSITY FOUNDATION INC] <brian.f.gore@nasa.gov>;

1 attachment

NASA TLX.PDF

Hello Ms. Campoe,

Thank you for your interest in our NASA TLX software; attached is a Software Usage Agreement for you to sign and return to us and we will send you the software.

Please enter your contact information on page two and the last page, sign and return to us.

Yes you may use the NASA-TLX paper version in your study funded by NSU.

Kim Chrestenson
Software Release Coordinator
Deltha/Critique

NASA Ames Research Center
M/S 202A-3
Moffett Field, CA 94035-1000
Phone (650) 604-5063
Fax (650) 604-7486
Ames Technology Partnerships Website
Software Release URL: https://nen.nasa.gov/web/sra

---

**From:** Kristi Campoe [mailto:kcampoe@knights.ucf.edu]
**Sent:** Monday, September 02, 2013 9:21 PM
**To:** Chrestenson, Kim L. (ARC-VP)[DELTHA-CRITIQUE]
**Subject:** Permission to use and access computerized NASA-TLX

Hello Ms. Chrestenson,

My name is Kristi Campoe. I am a PhD candidate at the University of Central Florida (UCF) in Orlando. I am also faculty with Nova Southeastern University (NSU) *College of Nursing*.

I am writing my UCF dissertation proposal in which I would like to study the impact of interruption frequency on registered nurses' performance, cognitive workload, and satisfaction after completing complex medical device programming tasks. With your permission, I would like to use the NASA-TLX computer/electronic version. I would like permission to use the NASA-TLX paper version in a related study funded by NSU.

If you require further information, please let me know. I look forward to your response.

Kristi R. Campoe, MSN, RN, CMSRN

## LIST OF REFERENCES


Altmann, E.M., Trafton, J.G. & Hambrick, D. Z. (2013). Momentary interruptions can derail the train of thought. *Journal of Experimental Psychology: General, 143*, 215-226.

Bagian, J.P. (2012). Healthcare and patient safety: The failure of traditional approaches and how human factors and ergonomics can and must help. *Human Factors and Ergonomics in Manufacturing & Service Industries 22*(1), 1-6.

Bangor, A., Kortum, P.T. & Miller, J.T. (2008). An empirical evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction, 24*(6), 574-594.

Bennett, J., Dawoud, D. & Maben, J. (2010). Effects of interruptions to nurses during medication administration. *Nursing Management - UK, 16*(9), 22-23.

Bevan, N. & Macleod, M. (1994). Usability measurement in context. *Behaviour & Information Technology, 13*(1-2), 132-145.

Biron, A. D., Loiselle, C. G. & Lavoie-Tremblay, M. (2009). Work interruptions and their contribution to medication administration errors: An evidence review. *Worldviews on Evidence-Based Nursing, 6*(2), 70-86.

Brace, N., Kemp, R., & Snelgar, R. (2006). *SPSS for psychologists* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Braun, S. (2005). *Usability for medical devices*. Paper presented at the 2005 IEEE Symposium on Product Safety Engineering, Schaumburg, IL.

Brixey, J., Zhang, J., Johnson, T., & Turley, J. (2009). Legibility of a volumetric infusion pump in a shock trauma ICU. *The Joint Commission Journal on Quality and Patient Safety, 35*(4), 229-235.

Campoe, K.R. (2013a). *Medical device usability analyses: An integrative review.* Paper presented at the 2013 International Symposium on Human Factors and Ergonomics in Health Care: Advancing the Cause, Baltimore, MD. Retrieved from http://hcs.sagepub.com/content/2/1/123.abstract

Campoe, K.R. (2013b). *Model of nurses' interaction with medical devices.* Unpublished manuscript. College of Nursing. University of Central Florida.

Campoe, K. R., Barnett, J. S., & Byers, J. F. (2012, Feb.). Human factors, nursing practice, and patient safety: An exemplar of the heuristic evaluation method. Paper presentation at the Nova Southeastern University Research Day, Ft. Lauderdale, FL.

Carayon, P., Hundt, A.S., & Wetterneck, T. (2010). Nurses' acceptance of Smart IV pump technology. *International Journal of Medical Informatics, 79*(6), 401-411.

Carayon, P., Wetterneck, T.B., Hundt, A.S., Ozkaynak, M., DeSilvey, J., Ludwig, B.,…& Rough, S. (2007). Evaluation of nurse interaction with bar code medication administration technology in the work environment. *Journal of Patient Safety, 3*(1), 34-42.

Carlton, G., & Blegen, M. A. (2006). Medication-related errors: A literature review of incidence and antecedents. *Annual Review of Nursing Research, 24*, 19-38.

Chan, A.J., Islam, M.K., Rosewall, T., Jaffray, D.A., Easty, A.C., & Cafazzo, J.A. (2012). Applying usability heuristics to radiotherapy systems. *Radiotherapy and Oncology, 102*(1), 142-147.

Chiu, C.C., Vicente, K.J., Buffo-Sequeira, I., Hamilton, R.M., & McCrindle, B.W. (2004). Usability assessment of pacemaker programmers. *Pacing and Clinical Electrophysiology: PACE, 27*(10), 1388-1398.

Crisp, C. C., Bandi, S., Kleeman, S. D., Oakley, S. H., Vaccaro, C. M., Estanol, M. V.,… & Pauls, R. N. (2012). Patient-controlled versus scheduled, nurse-administered analgesia following vaginal reconstructive surgery: A randomized trial. *American Journal of Obstetrics and Gynecology, 207*(5), 433.e1–6.

Czaja, S.J. (1997). System design and evaluation. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (2nd Ed.). New York, NY: John Wiley & Sons, Inc.

Dickinson, A., McCall, E., Twomey, B., & James, N. (2010). Paediatric nurses' understanding of the process and procedure of double-checking medications. *Journal of Clinical Nursing, 19*(5-6), 728-735.

Drews, F.A., & Kramer, H.S. (2012). Human–computer interaction design in health care, In P. Carayon (Ed.), *Handbook of Human Factors and Ergonomics in Health Care and Patient Safety* (pp. 265-277): Boca Raton, FL: CRC Press.

Dubey, S. K., & Rana, A. (2010). Analytical roadmap to usability definitions and decompositions. *International Journal of Engineering Science and Technology, 2*(9), 4723 - 4729.

Elo, S., & Kyngas, H. (2007). The qualitative content analysis process. *Journal of Advanced Nursing, 62*(1), 107-115.

Etchells, E., Bailey, C., Biason, R., DeSousa, S., Fowler, L., Johnson, K.,…& O'Neill, C. (2006). Human factors in action: Getting "pumped" at a nursing usability laboratory. *Healthcare Quarterly, 9*(Sp), 69-74.

Fairbanks, R.J., Caplan, S.H., Bishop, P.A., Marks, A.M., & Shah, M.N. (2007). Usability study of two common defibrillators reveals hazards. *Annals of Emergency Medicine, 50*(4), 424-432.

Folmer, E. & Bosch, J. (2004). Architecting for usability: A survey. *Journal of Systems and Software, 70*(1-2), 61-78.

Garmer, K., Liljegren, E., Osvalder, A.L., & Dahlman, S. (2002). Application of usability testing to the development of medical equipment. Usability testing of a frequently used infusion pump and a new user interface for an infusion pump developed with a human factors approach. *International Journal of Industrial Ergonomics, 29*(3), 145-159.

Garmer, K., Liljegren, E., Osvalder, A.L., & Dahlman, S. (2002). Arguing for the need of triangulation and iteration when designing medical equipment. *Journal of Clinical Monitoring and Computing, 17*(2), 105-114.

George, D. & Mallery, P. (2010). *SPSS for Windows step by step: A simple guide and reference, 18.0 update* (11th ed.). Boston, MA: Allyn and Bacon.

Ginsburg, G. (2005). Human factors engineering: A tool for medical device evaluation in hospital procurement decision-making. *Journal of Biomedical Informatics, 38*(3), 213-219.

Graham, M.J., Kubose, T.K., Jordan, D., Zhang, J., Johnson, T.R., & Patel, V.L. (2004). Heuristic evaluation of infusion pumps: Implications for patient safety in intensive care units. *International Journal of Medical Informatics, 73*(11-12), 771-779.

Grundgeiger, T., & Sanderson, P. (2009). Interruptions in healthcare: Theoretical views. *International Journal of Medical Informatics, 78*(5), 293-307.

Hall, L.M., Ferguson-Paré, M., Peter, E., White, D., Besner, J., Chisholm, A., & Hemingway, A. (2010). Going blank: factors contributing to interruptions to nurses' work and related outcome. *Journal of Nursing Management, 18*(8), 1040-1047.

Hart, S.G. (2006). NASA-Task Load Index (NASA-TLX) 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50*(9), 904-908.

Hart, S.G., & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload*. Amsterdam: North Holland Press.

Healey, A.N., Sevdalis, N., & Vincent, C.A. (2006). Measuring intra-operative interference from distraction and interruption observed in the operating theatre. *Ergonomics, 49*(5-6), 589-604.

Hicks, R., Sikirica, V., Nelson, W., Schein, J., & Cousins, D. (2008). Medication errors involving patient-controlled analgesia. *American Journal of Health-System Pharmacists., 65*, 429-440.

Hvannberg, E.T., Law, E.L.C., & Lárusdóttir, M.K. (2007). Heuristic evaluation: Comparing ways of finding and reporting usability problems. *Interacting with Computers, 19*(2), 225-240.

Hoonakker, P., Carayon, P., Gurses, A. P., Brown, R., Khunlertkit, A., McGuire, K., & Walker, J. M. (2011). Measuring workload of ICU nurses with a questionnaire survey: the NASA Task Load Index (TLX). *IIE Transactions on Healthcare Systems Engineering, 1*(2), 131-143.

Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies, 64*(2), 79-102.

Hudcova, J., McNicol, E. D., Quah, C. S., Lau, J., & Carr, D. B. (2011). Patient controlled opioid analgesia versus conventional opioid analgesia for postoperative pain. *Cochrane Database of Systematic Reviews* (3).

Institute for Safe Medication Practices. (2010). Proceedings from the ISMP Summit on the Use of Smart Infusion Pumps: GUIDELINES FOR SAFE IMPLEMENTATION AND USE. Retrieved from http://www.ismp.org/Tools/guidelines/smartpumps/printerVersion.pdf

International Organization of Standards [ISO]. (1998). ISO 9241-11 Guidance on usability. Retrieved from https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-1:v1:en

Jaspers, M.W. (2009). A comparison of usability methods for testing interactive health technologies: Methodological aspects and empirical evidence. *International Journal of Medical Informatics, 78*(5), 340-353.

Jennings, B.M., Sandelowski, M., & Mark, B. (2011). The nurse's medication day. *Qualitative Health Research, 21*(10), 1441-1451.

Kataoka, J., Sasaki, M., & Kanda, K. (2011). Effects of mental workload on nurses' visual behaviors during infusion pump operation. *Japan Journal of Nursing Science, 8*(1), 47-56.

Kjeldskov, J., Skov, M.B., & Stage, J. (2010). A longitudinal study of usability in health care: Does time heal? *International Journal of Medical Informatics, 79*(6), e135-e143.

Kushniruk, A., & Patel, V. (2004). Cognitive and usability engineering methods for the evaluation of clinical information systems. *Journal of Biomedical Informatics, 37*(1), 56-76.

Langdon, P., Persad, U., & John Clarkson, P. (2010). Developing a model of cognitive interaction for analytical inclusive design evaluation. *Interacting with Computers, 22*(6), 510-529.

Li, S.Y.W., Magrabi, F., & Coiera, E. (2012). A systematic review of the psychological literature on interruption and its patient safety implications. *Journal of the American Medical Informatics Association, 19*(1), 6-12.

Liljegren, E. (2006). Usability in a medical technology context assessment of methods for usability evaluation of medical equipment. *International Journal of Industrial Ergonomics, 36*(4), 345-352.

Lin, L., Isla, R., Doniz, K., Harkness, H., Vicente, K.J., & Doyle, D.J. (1998). Applying human factors to the design of medical equipment: Patient-controlled analgesia. *Journal of Clinical Monitoring and Computing, 14*(4), 253-263.

Lin, L., Vicente, K.J., & Doyle, D.J. (2001). Patient safety, potential adverse drug events, and medical device design: A human factors engineering approach. *Journal of Biomedical Informatics, 34*(4), 274-284.

Liu, Y., Tech, L., & Osvalder, A.L. (2004). Usability evaluation of a GUI prototype for a ventilator machine. *Journal of Clinical Monitoring and Computing, 18*(5), 365-372.

Marini, S.D., Hasman, A., Huijer, H.A., & Dimassi, H. (2010). Nurses' attitudes toward the use of the bar-coding medication administration system. *CIN: Computers, Informatics, Nursing, 28*(2), 112-123.

Martin, J.L., Norris, B.J., Murphy, E., & Crowe, J.A. (2008). Medical device development: The challenge for ergonomics. *Applied Ergonomics, 39*(3), 271-283.

Melynk, B. M., & Finout-Overholt, E. (2011). *Evidence-based practice in nursing and healthcare: A guide to best practice*. Philadelphia: Lippincott, Williams & Wilkins.

Meissner, B., Nelson, W., Hicks, R., Sikirica, V., Gagne, J., & Schein, J. (2009). The rate and costs attributable to intravenous patient-controlled analgesia errors. *Hospital Pharmacy, 44*(4), 312-324.

Nemeth, C., Nunnally, M., Bitan, Y., Nunnally, S., & Cook, R. I. (2009). Between choice and chance: The role of human factors in acute care equipment decisions. *Journal of Patient Safety, 5*(2), 114-121.

Nielsen, J. (1993). *Usability engineering*. San Diego, CA: Morgan Kaufmann.

Nielsen, J. (1994). *Enhancing the explanatory power of usability heuristics*. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence, Boston, MA.

Nielsen, J., & Mack, R. (1994). *Usability inspection methods*. New York, NY: Wiley.

Nunnally, M.E., & Bitan, Y. (2006). Time to get off this pig's back? The human factors aspects of the mismatch between device and real-world knowledge in the health care environment. *Journal of Patient Safety 2*(3), 124-131.

Obradovich, J.H., & Woods, D.D. (1996). Users as designers: how people cope with poor HCI design in computer-based medical devices. *Human Factors, 38*(4), 574-592.

Oxford Dictionaries. (2015). Attitude. Retrieved from http://www.oxforddictionaries.com/us/definition/american_english/attitude

Oxford Dictionaries. (2014). Reaction. Retrieved from http://www.oxforddictionaries.com/us/definition/american_english/reaction

Palese, A., Sartor, A., Costaperaria, G., & Bresadola, V. (2009). Interruptions during nurses' drug rounds in surgical wards: observational study. *Journal of Nursing Management, 17*(2), 185-192.

Palmer, E., Keebler, J., Fouquet, S., Lazzara, E., Simmons, J., & Chan, Y.R. (2013). *Effects of interruption type and interruption attitudes on doctors' and nurses' feelings of being "overwhelmed".* Paper presented at the 2013 International Symposium on Human Factors and Ergonomics in Health Care, Baltimore, MD.

Parasuraman, R. & Manzey, D.H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 52*(3), 381-410.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39*(2), 230-253.

Patel, V.L., & Currie, L.M. (2005). Clinical cognition and biomedical informatics: Issues of patient safety. *International Journal of Medical Informatics, 74*(11-12), 869-885.

Patel, V.L., & Kaufman, D.R. (1998). Medical informatics and the science of cognition. *Journal of the American Medical Informatics Association, 5*(6), 493-502.

Potter, P., Wolf, L., Boxerman, S., Grayson, D., Sledge, J., Dunagan, C., & Evanoff, B. (2005). Understanding the cognitive work of nursing in the acute care environment. *Journal of Nursing Administration, 35*(7/8), 327-335.

Redding, D.A., & Robinson, S. (2009). Interruptions and geographic challenges to nurses' cognitive workload. *Journal of Nursing Care Quality, 24*(3), 194-200.

Rice, S. (2009). Examining single- and multiple-process theories of trust in automation. *Journal of General Psychology, 136*(3), 303-319.

Rivera-Rodriguez, A.J., & Karsh, B.T. (2010). Interruptions and distractions in healthcare: Review and reappraisal. *Quality & Safety in Health Care, 19*(4), 304-312.

Rogers, W.A., Mykitshyn, A.L., Campbell, R.H., & Fisk, A.D. (2001). Analysis of a "simple" medical device *Ergonomics in Design: The Quarterly of Human Factors Applications, 9*(1), 6-14.

Schein, J. R., Hicks, R. W., Nelson, W. W., Sikirica, V., & Doyle, D. J. (2009). Patient-controlled analgesia-related medication errors in the postoperative period: causes and prevention. *Drug Safety: An International Journal of Medical Toxicology and Drug Experience, 32*(7), 549-559.

Schraagen, J. M., & Verhoeven, F. (2013). Methods for studying medical device technology and practitioner cognition: The case of user-interface issues with infusion pumps. *Journal of Biomedical Informatics, 46*(1), 181-195.

Shackel, B. (1991). Usability - context, framework, definition, design and evaluation. In B. Shackel & S. J. Richardson (Eds.), *Human factors for informatics usability*. Cambridge, MA: Cambridge University Press.

Shackel, B. (2009). Usability – Context, framework, definition, design and evaluation. *Interacting with Computers, 21*(5–6), 339-346.

Shah, S.G., & Robinson, I. (2007). Benefits of and barriers to involving users in medical device technology development and evaluation. *International Journal of Technology Assessment in Health Care, 23*(1), 131-137.

Shneiderman, B., & Plaisant, C. (2010). *Designing the user interface: Strategies for effective human-computer interaction*. Boston, MA: Addison-Wesley.

Sörqvist, P. (2010). The role of working memory capacity in auditory distraction: A review. *Noise & Health, 12*(49), 217-224.

Tabachnick, B. G., & Fidell, L. S. (2012). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson.

Thyvalikakath, T.P., Monaco, V., Thambuganipalle, H., & Schleyer, T. (2009). Comparative study of heuristic evaluation and usability testing methods. *Studies in Health Technology & Informatics, 143*, 322-327.

Trbovich, P., Prakash, V., Stewart, J., Trip, K., & Savage, P. (2010). Interruptions during the delivery of high-risk medications. *Journal of Nursing Administration, 40*(5), 211-218.

Treiber, L.A., & Jones, J.H. (2010). Devastatingly human: An analysis of registered nurses' medication error accounts. *Qualitative Health Research, 20*(10), 1327-1342.

Trivedi, M.C., & Akheela Khanum, M. (2012). Role of context in usability evaluations: A review. *Advanced Computing: An International Journal (ACIJ), 3*(2).

Trbovich, P.L., Pinkney, S., Cafazzo, J.A., & Easty, A.C. (2010). The impact of traditional and smart pump infusion technology on nurse medication administration performance in a simulated inpatient unit. *Quality and Safety in Health Care, 19*(5), 430-434.

Trbovich, P., Prakash, V., Stewart, J., Trip, K., & Savage, P. (2010). Interruptions during the delivery of high-risk medications. *Journal of Nursing Administration, 40*(5), 211-218.

Turley, J.P., Johnson, T.R., Smith, D.P., Zhang, J., & Brixey, J.J. (2006). Operating manual-based usability evaluation of medical devices: An effective patient safety screening method. *Joint Commission Journal on Quality And Patient Safety / Joint Commission Resources, 32*(4), 214-220.

U. S. Food and Drug Administration. (2011). *Draft guidance for industry and Food and Drug administration staff - Applying human factors and usability engineering to optimize medical device design*. Rockville, MD: Center for Devices and Radiological Health-Office of Device Evaluation. Retrieved from http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm259748.htm

U.S. Food and Drug Administration. (2010). *About the FDA*. Retrieved from http://www.fda.gov/AboutFDA/default.htm

U.S. Food and Drug Administration. (2010). *Medical device home use initiative*. Retrieved from: www.fda.gov/

U.S. Food and Drug Administration. (2000). *Medical devices use-safety: Incorporating human factors engineering into risk management*. Washington, D. C.: U.S. Department of Health and Human Services. Retrieved from www.fda.gov/.

U.S. Food and Drug Administration. (2010). *Total product life cycle: Infusion pump premarket notification draft guidance*. Rockville, MD: Retrieved from http://www.fda.gov/

U.S. Food and Drug Administration. (2010). *White paper: Infusion pump improvement project*. U.S. Department of Health and Human Resources Retrieved from http://www.fda.gov/

Westbrook, J.I., Coiera, E., Dunsmuir, W.T.M., Brown, B.M., Kelk, N., Paoloni, R., & Tran, C. (2010). The impact of interruptions on clinical task completion. *Quality & Safety in Health Care, 19*(4), 284-289.

Westbrook, J.I., Woods, A., Rob, M.I., Dunsmuir, W., & Day, R.O. (2010). Association of interruptions with an increased risk and severity of medication administration errors. *Archives of Internal Medicine, 170*(8), 683-690.

Wolf, L.D., Potter, P., Sledge, J.A., Boxerman, S.B., Grayson, D., & Evanoff, B. (2006). Describing nurses' work: Combining quantitative and qualitative analysis. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 48*(1), 5-14.

Wickens, C.D. (1992). *Engineering psychology and human performance*: New York, NY: Harper Collins.

Wickens, C.D., & Carswell, C. M. (1997). Information processing. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (2nd ed.). New York, NY: John Wiley & Sons, Inc.

Wickens, C., & Hollands, J. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, N.J.: Prentice Hall.

Wickens, C., & McCarley, J. (2008). *Applied attention theory*. Boca Raton, FL: CRC Press Taylor & Francis Group.

Wicklund, M., Kendler, J., & Strochlic, A. Y. (2011). *Usability testing of medical devices*. New York: CRC Press Taylor & Francis Group.

Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. *Journal of Advanced Nursing, 52*(5), 546-553.

Zhang, J., Johnson, T.R., Patel, V.L., Paige, D.L., & Kubose, T. (2003). Using usability heuristics to evaluate patient safety of medical devices. *Journal of Biomedical Informatics, 36*(1-2), 23-30.

Zuzelo, P.R., Gettis, C., Hansell, A.W., & Thomas, L. (2008). Describing the influence of technologies on registered nurses' work. *Clinical Nurse Specialist: The Journal for Advanced Nursing Practice, 22*(3), 132-142.